

PSYCHOLOGICAL REVIEW

GEORGE MANDLER, EDITOR

JEAN M. MANDLER, ASSOCIATE EDITOR

UNIVERSITY OF CALIFORNIA, SAN DIEGO

Vol. 82, No. 1

January 1975

Articles

Consequences of Commitment to and Disengagement
from Incentives **Eric Klinger** 1

Toward a Psychological Theory of Alienation
Daniel Stokols 26

Sentence Comprehension: A Psycholinguistic Process-
ing Model of Verification
Patricia A. Carpenter and Marcel Adam Just 45

Birth Order and Intellectual Development
R. B. Zajonc and Gregory B. Markus 74

Editorial Consultants

The editor gratefully acknowledges the efforts of the following consultants who helped review manuscripts for this issue. They are part of a large group of psychologists who have kindly agreed to assist in the editing of manuscripts for *Psychological Review*.

J. W. ATKINSON	CHARLES CRAWFORD	DANIEL KATZ	EDWARD E. SMITH
LILLIAN BELMONT	JOAN E. FOLEY	SAMUEL S. KOMORITA	HARRY G. SPERLING
DARYL BEM	MARSHALL HAITH	BIBB LATANÉ	CHARLES F. STOMEYER III
ROBERT BOICE	O. J. HARVEY	LEAH LIGHT	EARL THOMAS
ROBERT C. BOLLES	JOHN R. HAYES	DAVID PREMACK	HOBEN THOMAS
KENNETH S. BOWERS	JERRY HIRSCH	DAVID E. RUMELHART	HANS WALLACH
BRUNO BREITMEYER	LEO M. HURVICH	KURT SCHLESINGER	WAYNE A. WICKELGREN
DORWIN CARTWRIGHT	FRANCIS W. IRWIN	CARMI SCHOOLER	

The *Psychological Review* publishes articles that make theoretical contributions to any area of scientific psychology. Preference is given to papers that advance theory rather than review it, and to statements that are specifically theoretical rather than programmatic. Surveys of the literature, problems of method and design, and papers that primarily report empirical findings are ordinarily not appropriate. Letters not exceeding 250 words on topics of theoretical interest should be submitted in duplicate.

Preparation of Manuscripts: Instructions for the preparation of manuscripts appear in the *Publication Manual of the American Psychological Association* (2nd ed.). Submit three copies; one copy must be the original typed copy. All parts of the manuscript must be double spaced. In order to facilitate blind reviewing, the author's name should not appear on the manuscript or abstract, but each copy of the manuscript should include a separate title page showing the title and the author's name and affiliation. Footnotes that identify the author should be typed on separate pages. Figures should be submitted as glossy prints ready for reproduction; duplicate figures may be photocopies. Metric equivalents must be given for all nonmetric measurements. Authors should keep a copy of the manuscript to guard against loss.

Abstracts: Manuscripts must be accompanied by an abstract of 100 to 175 words typed on a separate sheet of paper. Detailed instructions for preparation of abstracts appear in the *Publication Manual* or may be obtained from the Editor.

Manuscripts: Submit manuscripts in triplicate to the Editor, GEORGE MANDLER, Department of Psychology, University of California, San Diego, La Jolla, California 92037 according to instructions provided above.

Copyright and Permissions: All rights reserved. Written permission must be obtained from the American Psychological Association for copying or reprinting text of more than 500 words, tables, or figures. Permissions are normally granted contingent upon like permission of the author, inclusion of the APA copyright notice on the first page of reproduced material, and payment of a fee of \$10 per page of APA material. Permission fees are waived for authors who wish to reproduce their own articles. Address requests for permission to the Permissions Office, American Psychological Association, 1200 Seventeenth Street, N.W., Washington, D. C. 20036.

Subscriptions: Subscriptions are available on a calendar year basis only (January through December). Nonmember rates for 1975: \$15 domestic, \$16 foreign, \$3 single issue. APA member rate: \$6. Write to Subscription Section, APA.

Reprints: Authors may order reprints of their articles from the printer when they receive galley proofs.

Back Issues and Back Volumes: For information regarding back issues or back volumes write to Order Department, APA.

Microform Editions: For information regarding microform editions write to any of the following: Johnson Associates, Inc., P.O. Box 1017, Greenwich, Connecticut 06830; University Microfilms, Ann Arbor, Michigan 48106; or Princeton Microfilms, Princeton, New Jersey 08540.

Change of Address: Send change of address notice and a recent mailing label to the attention of the Subscription Section, APA, 30 days prior to the actual change of address. Undelivered copies resulting from address changes will not be replaced; second-class mail is not forwarded unless subscribers notify the local post office in writing that they will guarantee second-class forwarding postage.

Published bimonthly (beginning in January) in one volume per year, six issues, by the American Psychological Association, Inc., 1200 Seventeenth Street, N. W., Washington, D. C. 20036. Printed in the U.S.A. Second-class postage paid at Washington, D. C., and additional mailing offices.

HAROLD P. VAN COTT
Managing Editor

ANITA DEVIVO
Executive Editor and
General Manager

DEBORAH S. FELDER
Supervising Technical
Editor

MARY GRIMBALL
Technical Editor

LYNN D. SPIER
Acting Manager, Publications
Business Services

ANNE REDMAN
Subscription Manager

PSYCHOLOGICAL REVIEW

Copyright © 1975 by the American Psychological Association, Inc.

VOL. 82, No. 1

JANUARY 1975

Consequences of Commitment to and Disengagement from Incentives

Eric Klinger
University of Minnesota, Morris

Behavior and experience are organized around the enjoyment and pursuit of incentives. During the time that an incentive is behaviorally salient, an organism is especially responsive to incentive-related cues. This sustained sensitivity requires postulating a continuing state (denoted by a construct, *current concern*) with a definite onset (*commitment*) and offset (*consummation* or *disengagement*). Disengagement follows frustration, accompanies the behavioral process of extinction, and involves an *incentive-disengagement cycle* of invigoration, aggression, depression, and recovery. Depression is thus a normal part of disengagement that may be either adaptive or maladaptive for the individual but is probably adaptive for the species. The theory offers implications for motivation; etiology, symptomatology, and treatment of depression; drug use; and other social problem areas.

This article proposes a theoretical framework for accommodating psychological aspects of incentive relationships. The theory is primarily integrative rather than reductive, describing consistent relationships among various coordinate classes of observations and concepts rather than proposing specific mechanisms for them at another level of analysis. Nevertheless, the theory does suggest something about the properties of underlying behavioral and neurohumoral

structures, provides analyses of several psychological problem areas, and generates specific predictions concerning motivational, emotional, and cognitive functioning.

CENTRAL STATUS OF INCENTIVE IN MOTIVATIONAL THEORY

The theoretical development takes as its point of departure an organism's relationship to its incentives, defined here conventionally as objects or events that attract an organism or repel it. The concept of incentive has steadily gained prominence in motivational theory for the past three decades at the expense of the two other concepts that have dominated the field during this time: drive to describe internal motivational states and reinforcement (in the operant sense) to describe an organism's commerce with its environment. Both of these concepts have encountered serious theoretical difficulties.

The problems of drive stem principally from the growing evidence that whereas drive states as conventionally defined are

Various phases of the research described in this article were supported by National Science Foundation Grant GS-2735, National Institute of Mental Health Grant 1-R01-MH24804-01, and grants from the University of Minnesota Graduate School. I thank Roxanne M. Anderson, John W. Atkinson, Steven G. Barta, Charles G. Costello, Donald W. Fiske, Mariam Frenier, Rachel E. Froiland, George B. Flamer, Clifton W. Gray, Heinz Heckhausen, Ernest D. Kemble, Donald Leavitt, Thomas Mahoney, Richard T. Santee, Deborah A. Smith, and Warner Wilson for helpful comments on earlier drafts of this article.

Requests for reprints should be sent to Eric Klinger, Division of Social Sciences, University of Minnesota, Morris, Morris, Minnesota 56267.

neither necessary nor sufficient to launch sequences of goal-directed behavior, expectations that the behavior will lead to an incentive are both necessary and sufficient (Bindra, 1968, 1974; Bolles, 1967; Klinger, 1971). Dissatisfaction with the drive concept is thus widening (see also Appley, 1970; Atkinson, 1964; Cofer & Appley, 1964). Although drive states affect activity levels (Bindra, 1968), determine the attractiveness of at least certain kinds of incentives (such as food) at the point of consummation, and generate internal stimuli, by themselves they do not instigate instrumental behavior or the content of fantasy and dreams. Vigorous behavior can begin in the virtual absence of detectable drive, and fairly stable behavioral sequences can continue despite drastic fluctuations in the degree and configuration of drive states. It is therefore necessary to find a concept other than drive to tie together the sometimes elaborate long-term sequences of goal-directed behavior.

The concept of operant reinforcement has fallen on potentially even harsher times. It has been cogently argued that the concept is theoretically either unnecessary or erroneous (Atkinson & Wickens, 1971; Bolles, 1972; Estes, 1969, 1971; Tapp, 1969; Walker, 1969). While nobody denies that the *operation* of reinforcement (that is, dispensing rewards) often works to modify behavior, the occasions on which it fails and, perhaps more important, the manner in which it often succeeds disconfirm the theoretical *process* of reinforcement. Although there is still recent support for the notion that reinforcement incrementally strengthens stimulus-response (S-R) bonds (e.g., Berlyne, 1967; Perkins, 1968) and hence produces S-R learning, there is increasing evidence that reinforcement does not affect S-R bonds at all (Atkinson & Wickens, 1971; Estes, 1971; Levine, 1971; Logan, 1971; Nuttin & Greenwald, 1968) but only certain other processes that ultimately affect performance. Rewards and punishments provide feedback concerning the effectiveness of the organism's behavior, attract attention, determine what gets stored in short-term memory or transferred to long-term

memory, and determine where the organism will allocate its resources of time and effort. Since the effects of rewards and punishments can thus be attributed to a number of processes and events that have independent theoretical stature, these effects need not be attributed to any separate process called "reinforcement."

By vesting in incentive-related processes certain crucial functions that have previously been attributed to drive, and by limiting the status of reinforcement to operations that manipulate incentives, the argument places a heavy explanatory burden on the concept of incentive and hence raises pressing questions about how to conceptualize the means by which incentives can exert such far-reaching effects. This article presents an approach toward such a theory of incentive-related mechanisms by examining and conceptualizing the consequences to an organism of its organization around incentives.

CONSEQUENCES OF COMMITMENTS TO INCENTIVES

Two Constructs: Current Concern and Commitment

In relatively simple species, pursuing a goal can more accurately be described as a two-fold process of, first, engaging in actions that increase the likelihood of encountering the goal object and, second, being biochemically and behaviorally receptive to consummating the goal object. Higher species, of course, retain those characteristics, but their goal striving depends less on immediate cues and is therefore more persistent. A mosquito wafted out of sensory contact with a warm body is unlikely to retrace its steps in order to take up where it left off, but a dog that has lost sight of a thrown stick may search for a while, and a person trying to create an experimental result may go on doggedly for years searching for the right combination of conditions.

An organism that can keep on pursuing a goal despite many changes in drive states and environmental cues, even in the face of repeated obstacles, must possess a continuing neural process that somehow maintains

the goal as a controlling influence on behavior. In the terminology of this article, the hypothetical event that marks the onset of this process is called the organism's becoming "committed"¹ to that particular goal. The "state" of having become committed to a particular goal is here called a "current concern" about achieving the goal. By this definition, an organism may be in the throes of several or even numerous current concerns (and hence numerous states) at a time. For instance, a student may be committed to pursuing a degree, completing certain courses, maintaining or improving his relationship with a close friend, finding a part-time job, dinner, skiing at year's end, and so on. Each of the ends to which he is committed constitutes an incentive (because he is attracted to it) and a goal (because he is committed to striving for it), and about each he has a current concern (the construct that represents his state of being committed to that goal).

The concept of current concern is, at this time in our theory building, a provisional, heuristic, nonmentalistic construct devised to label a discrete state: the state of the organism between the time of commitment and the time of consummation or disengagement. The influence that states of concern exert on behavior depends on a number of variables that can be investigated empirically, but these are not part of the concern construct itself. Hence, current concern is unlike Atkinson and Birch's (1970) "action tendency." After commitment its existence does not depend on arousal by cues, and it has no dimension of "strength" that depends on situational factors, although the behavior

that arises during the state is, of course, affected by an array of moderator variables. Thus, current concern is theoretically compatible with such other concepts as action tendency, but it is formally different and it performs different functions theoretically and heuristically.

It may be helpful in clarifying the concept of current concern to contrast it with some other concepts. It resembles Lewin's (1928) concept of "quasineed," but it does not share the distinction between quasineeds and "genuine needs." It implies narrower classes of satisfiers than "need" (Murray, 1938) or "motive" (McClelland, Atkinson, Clark, & Lowell, 1953; e.g., commitment to setting a new sales record rather than to achievement). It differs from "drive" and older uses of "need" in eliminating connotations of both tissue deficits and of activity in neural circuits commonly identified with consummatory behavior or reward sites. Finally, it differs from "central motive state" (Bindra, 1974; Morgan, 1943) in referring in each usage to a potentially enduring state rather than to one elicited by situational cues. Current concerns are conceived here to persist even during intervals when the organism is engaged in behavior irrelevant to the concern, although they continue to influence behavior in more subtle ways.

Most previous research approaches have proceeded without a concept of commitment as an onset event warranting investigation in its own right, and apparently none has provided either evidence concerning the hypothetical process of commitment as such, apart from its sequelae, or independent measures of it. At this point, therefore, the existence of a commitment process must be inferred² as a logically necessary antecedent

¹ The construct "commitment" is formally somewhat similar to Lewin's (1928) concept of *Vornahme*. It may also seem to resemble Irwin's (1971) concept of "intention." These terms, however, imply a property of being conscious which I do not wish to stipulate. Furthermore, Irwin defines intention as involving a three-fold correspondence among a subject's preference for a specified outcome, an expectancy that a certain act will lead to that outcome, and a second expectancy that an alternative act would lead to a nonpreferred outcome. This definition lacks the temporal specificity of the commitment construct as the onset of a state. It seems best to distinguish the two concepts by using different terms.

² Because there are at present no independent measures of commitment, operational definitions that an organism has become committed to something must take one of two forms. First, one can observe some of the response classes that presumably follow commitment such as detour behavior, preoccupation with an incentive, etc. This is not necessarily a circular procedure, since under the present theory knowing some of these behaviors permits prediction of others. A second form of operational definition is to arrange conditions that

target word over a relatively brief time span, conforms with operations designed to produce effects of "instructional sets." However, existing evidence suggests that people can remain sensitized beyond the range of such a time frame.

For instance, both waking and sleeping subjects are continuously sensitive to their own names. In another dichotic listening experiment, subjects tended to notice instructions on the nonattended channel when these were preceded by the subject's name (Moray, 1959). This result can be taken to indicate shifts of attention to concern-related cues if one assumes that subjects have learned that mention of their names is usually relevant to one or another of their current concerns. These results cannot be explained as effects of instructional sets and could not have involved information in short-term storage.

It could be argued that one's own name is so powerfully conditioned to arouse emotions that it is too nonspecific a stimulus to use for assessing the accompaniments of current concerns, but other evidence from sleeping subjects suggests that the Moray results need not be explained solely in terms of conditioned emotional arousal. Subjects who were put to sleep in a laboratory with instructions to clench their fists whenever they heard an arbitrarily selected name (not their own) managed to obey this instruction on one eighth of the playings, whereas they very rarely clenched their fists in response to 54 control names (Oswald, Taylor, & Treisman, 1960).

One important issue raised by the current concerns approach to data such as these is whether organisms remain responsive to concern-related cues throughout the interval between commitment and consummation (or disengagement) or whether they become responsive only after they have been aroused on a particular occasion. Although it seems likely that arousing a subject's emotions or expectancies regarding an incentive will increase responsivity toward cues related to it (Klinger, 1973), the evidence suggests that not all of the increased responsivity to concern-related cues can be explained in this way. Thus, for instance,

subjects' tendency in the experiment by Oswald et al. (1960) to clench their fists when they heard assigned names other than their own would have required arousal over such a long time period and under such unfavorable conditions (sleep) as to confirm the need for the kind of construct represented by the concept of current concern. Other kinds of behavior commonly observed outside the laboratory, such as a sleeping mother's specific sensitivity to her young infant's noises or a sleeper's ability to wake up at approximately a preset time (Tart, 1970), also suggest the need to postulate a continuing neural process that is not dependent on the kind of short-term arousal produced by situational cues or emotional responses.

Some unpublished data that bear on these questions are beginning to emerge with the use of thought-sampling methods for the study of mental content. Subjects listened dichotically to tape-recorded narrations of fiction, segments of which had been rewritten to allude unobtrusively to individual subjects' current concerns or to matters believed not to be concerns. Subjects spent more time attending to and were more likely to recall concern-related passages than nonconcern-related passages. Second, the thought contents that they reported on these occasions were more likely to be thematically related to the preceding passages if the passages had been recalled than if they had not and if the passages were concern related. Since the subjects had not received any specific procedures to sensitize them to their concerns, which had been inferred from interviews and questionnaires administered on previous days, their differential sensitivity to concern-related stimuli cannot be explained on grounds of situational arousal.

The evidence that subjects are especially sensitized or attentive to certain concern-related cues suggests that the preattentive processes (Neisser, 1967) that control shifts in attention are attuned to material related to a person's current concerns. This, then, may serve as the mechanism that accounts for the preponderance of concern-related content in fantasy and dreams.

Effects on Perception and/or Recall

There are by now many experiments on the effects of motivation and instructions on perception or perception-related responses (e.g., Haber, 1966; Jenkin, 1957; Postman, 1953). Although this research has demonstrated a variety of means for reliably influencing perception-related responses, it has so far not resolved the question of precisely what is affected: the perceptual experience itself, response biases, or the way in which the subject codes and organizes his memory of the percept (Haber, 1966). Information-processing analyses suggest a number of mutually compatible loci of influence (Erdelyi, 1974): search strategies, receptor modifications, encoding strategies, decisions regarding transfer to short- or long-term memory, and selection of overt responses. Most of these components are overt or covert instrumental acts and as such are clearly under the control of incentive expectancies. There remains the possibility, of course, that perception and recall are also enhanced motivationally in other ways.

The existing evidence is insufficient to pinpoint the specific processing steps in perceptual reporting that are affected by motivational variables. Nevertheless, it can at least be concluded provisionally that stimuli related to a subject's current concerns seem more prominent than other stimuli (Jenkin, 1957). Thus, for instance, subjects report need-related objects to look larger (Bruner & Goodman, 1947; Lambert, Solomon, & Watson, 1949) and brighter (Gilchrist & Nesberg, 1952) than they are, and attractive or unattractive male faces seem more familiar than neutral faces (Gerard, Green, Hoyt, & Conolley, 1972). Subjects are also likely to interpret ambiguous or very briefly exposed stimuli as representing objects that are related to current needs (Atkinson & McClelland, 1948; McClelland & Atkinson, 1948).

Results such as these were once interpreted as effects of drive. However, they can probably be better interpreted to reflect nondrive factors such as incentive expectancies. Saugstad (1966), for instance, concluded from a review of hunger effects on

perceptual and cognitive processes that "the important condition may not be hours of deprivation, but the expectancy of the Ss as to when they may next receive food" (p. 80).

Integrative Discussion

Up to this point, the theory as described accomplishes basically three kinds of things.

1. It establishes a set of concepts that refer to a hypothetical state (current concern) with a starting point (commitment) and end point (consummation or disengagement).

2. The concept of current concern constitutes a conceptual focus for the inter-relatedness of specified response classes (actions, mental content, preattentive processes, and perception) that might otherwise seem quite diverse, and its use argues that with respect to such response classes the organism is organized around its current concerns.

3. The theory pinpoints important motivational issues and facilitates analysis of a number of important problem areas. Each new concept raises fundamental questions concerning its parameters and the characteristics of the underlying process. (a) The underlying process presumably persists from commitment forward. Terminating it short of goal attainment requires an active countervailing process. It incorporates semantic codes for specific incentive-related cues and cue classes and maintains these in long-term storage. It contains comparator mechanisms for evaluating response feedback and determining goal attainment. (In the latter respects, it operates similarly to a "plan" [Miller, Galanter, & Pribram, 1960].) (b) The finding that people are especially responsive to concern-related cues (including, presumably, the concern-related cues emitted in the person's own previous thought) helps to explain the previously described effects of concern states on mentation (Klinger, 1971), since the increased responsiveness to concern-related cues would itself increase responses to those cues, which are, of course, likely to be concern related. Applying these concepts in this way generates further specific predictions about the effect of current concerns on the effectiveness of

creative thought and on the efficiency of problem-solving thought (fluency, distractibility, effortfulness) in general. (c) The basic notion that organisms are behaviorally (including ideationally and affectively) organized around incentives immediately raises the question as to what happens—to current concerns, to mental content, and to affect as well as to action—when a person finds himself in a situation barren of incentives. (d) The theory raises empirical questions as to the variables that moderate the influence of concerns on action, attention, thought, and affect. Preliminary analyses suggest determining variables such as those already suggested by Expectancy $\times$ Value theories as well as others such as the amount of time before something must be done and previous investments of time and effort. (e) There are basic questions regarding the nature of the commitment process, such as whether it is in fact a discrete event or whether it constitutes a continuous variable. (f) A further important issue involves the degree to which a state of concern is specific with respect to the particular instrumental acts, particular time relationships, and other attributes of goal attainment, and how these specifications are incorporated.

CONSEQUENCES OF DISENGAGEMENT FROM INCENTIVES

The conceptualization described thus far has dealt with the consequences of commitment as though the value of incentives remains constant, and it has said nothing concerning interferences with goal attainment. In fact, however, incentive values do fluctuate. They are most likely subject to habituation and to shifting adaptation levels. The incentive, once obtained, may not come up to expectation. Its value may vanish with a goal to which it was chained (Costello, 1972). Unforeseen difficulties in pursuing an incentive appear for a time to increase its value. Depression decreases many incentive values.

Although it seems likely that all lapses from the expected accessibility, timeliness, and quality of an incentive affect its value, the discussion that follows will, for the sake of simplicity, focus on the consequences of

complete failure, leading to the nullification of commitment. Given the powerful organizational properties possessed by states of current concern, it seemed likely that, when goals are unattainable or too costly and must therefore be abandoned without consummation, ending the process set in motion by a commitment should itself be a rather involved process. Furthermore, deductions from the theory suggested that dissolving an important concern should tend to produce effects such as apathy, reduced instrumental striving, loss of concentration, and increased preoccupation with momentary cues, a pattern that is recognizably similar to depression.

Fortunately, there is a large amount of evidence that can be brought to bear on this process since blocked goal striving is the operational definition of frustration and, under many circumstances, of extinction. It leads to a variety of consequences besides a cessation of operant activity—invigoration, primitivization, aggression, and depression. The purpose of this section is to present an integrated description of these consequences, showing them to be not contradictory but rather different phases of an orderly process of ending commitment to an incentive, a process of *disengagement* involving an *incentive-disengagement cycle*.

For purposes of exposition, the phenomena of the incentive-disengagement cycle are described in five categories or phases. Nevertheless, the cycle is intended to be depicted as continuous rather than as a sequence of discrete stages. Furthermore, there are circumstances in which the invigoration or aggression phases may be attenuated or absent during disengagement. Finally, a person may be in various phases of several different cycles simultaneously, each involving a different incentive, with the contemporaneous phases of the several cycles summing or interacting.

Invigoration

After an organism has launched an instrumental sequence, it responds to an obstacle with increased vigor, taking the form of more powerful motions, more rapid responses, more intense concentration, and so

forth. A form of the invigoration effect is well known under Amsel's label of "Frustration Effect" (Amsel, 1958, 1962).⁴ That is, animals run faster in a second runway after nonreward than after reward in a first runway. When disappointed animals are put into an open field right after nonreward instead of into a second runway, their response vigor also increases (Gallup & Altomari, 1969). Retardates rock more after nonreward (Forehand & Baumeister, 1971), and children respond more forcefully (Gilbert, 1969; Holton, 1961; Ryan & Watson, 1968), especially if they are stopped just short of (rather than far from) achieving a goal (Holton, 1961). The frustration effect is independent of whether Reward 1, which was omitted, resembles Reward 2 (Berger, 1969), the amount of time required to get the reward (Logan, 1968), or whether the animal's responses are thwarted after nonreward (DiLollo, Davidson, Hammond, & Donovan, 1968). Re-

tardates and children react more vigorously to frustration in accordance with the number of times the response had been rewarded prior to nonreward (Forehand & Baumeister, 1971; Holton, 1961). Mandler (1964) found that during extinction of an overlearned response, rats "exhibit both exploratory behavior . . . and repeated attempts to perform the interrupted sequence" (p. 196). People express more need for achievement in Thematic Apperception Test stories written right after failing at a task than after succeeding (McClelland et al., 1953). Although it is still possible that invigoration effects would disappear if the organism could be certain that its chances of goal attainment had become nil, an invigoration effect clearly has been demonstrated in widely varying investigations.

Frustration is reflected not only in behavioral invigoration but also in the value of the blocked incentive. When people are frustrated they come to regard the incentive as more attractive than before, judging from ratings (Brehm, 1972; Mischel & Masters, 1966) and from looking time, overestimation of size, and favorable statements about syllables that were associated with it (Knott, Nunnally, & Duchnowski, 1967). There is seemingly no evidence concerning what happens to the value of other incentives while the value of the blocked incentive rises. However, adaptation level theory (Appley, 1971) and incentive contrast phenomena (Adamson, 1971; Fowler, 1971) suggest that if one value rises sharply the other values should weaken. It therefore seems likely that during the invigoration and aggression phases not only does the blocked incentive become more attractive but also others become less so: the loss of "perspective" often noted in people who become over-involved in frustrated undertakings.

Aggression

There is little need to review here the extensive literature on the frustration-aggression effect (Dembo, 1931; Dollard et al., 1939). With appropriate qualifications (Berkowitz, 1962), frustration commonly produces aggression, even though other events may also sometimes produce aggres-

⁴ Staddon (1970, 1972) has presented an incisive analysis of weaknesses in the experimental evidence for an energizing frustration effect. Staddon proposes an alternative explanation, that the frustration effect reflects not enhanced responding after nonreward but inhibited responding after reward. However, his analysis does not seem to invalidate the argument presented here. Using an operant (key-pecking) design, he and some colleagues have shown that omitting reward for key-pecks at the end of fixed intervals increases response rate (that is, eliminates "scaloping") during the following intervals. However, when reward is contingent on *not* pecking (so as to produce an intrainterval negative goal gradient), omitting reward reduces pecking (that is, increases not pecking) on the following intervals. This would seem consistent with the notion advanced here that the first effect of frustration is to invigorate operant acts that have been learned to obtain reward. Staddon denies this explanation on grounds that when nonreward is accompanied by some of the cues (in his case, blackouts) that had accompanied reward, the following operant behavior is enhanced less than after nonreward unaccompanied by those reward-related cues. However, it can be assumed that reward-related cues would have acquired secondary incentive value, and would therefore (at least initially) dampen the extent of affective frustration. Results described below (Klinger, Barta, & Kemble, 1974) support a response invigoration interpretation and seem unaccountable in terms of Staddon's reward-dependent inhibition explanation.

sion and even though frustration need not always do so. With continued frustration responses become more stereotyped and more primitive until at last they often may be recognized as destructively aggressive. After frustration, children's play responses become more primitive (Barker, Dembo, & Lewin, 1941) as do adults' responses in tests of originality and divergent thinking (Hinton, 1968). Extinction in pigeons (Azrin, Hutchinson, & Hake, 1966) and in people (Kelly & Hake, 1970) and non-reward in rats (Gallup, 1965) have specifically been shown to produce eventual aggression. Even when subjects report no anger following failure, their blood pressure rises (Gentry, 1970). When a personal catastrophe sends someone into a reactive depression, it is very likely to send him first into disbelief and angry protest, as in blindness (Fitzgerald, 1970) and separation of human children (Bowlby, 1973) and infant monkeys (McKinney, Suomi, & Harlow, 1971) from their parents.

Downswing into Depression

At some point during sustained unrelieved frustration, organisms begin to give up. The available evidence suggests that failure, loss, and disappointment (in the objective sense) generally lead to depression-like responses that range along a continuum from mild disappointment (in the subjective sense) to severe depression.

The downswing into depression is probably not a matter of the person "turning" his anger "inward." In one study of this issue, clinically depressed patients expressed more resentment than nonpatients, and both their resentment and their verbal hostility decreased as the depression lifted (Friedman, 1970). When not depressed, the depression-prone members of this sample responded with less verbal hostility than the matched controls. College student subjects who provided daily self-ratings of mood yielded positive correlations between depression and anger both within subjects over time and across subjects (Wessman & Ricks, 1966). Thus, it seems clear that depression is accompanied by heightened hostility.

There are several lines of evidence to support the notion of a frustration-depression sequence. After subjects are experimentally induced to fail, they describe themselves as more depressed (Ellner, 1970; Lish, 1969), and they express less self-esteem than after successful completion of a task (Lish, 1969). If failure is accompanied by much frustration, the Thematic Apperception Test stories subjects write immediately thereafter express reduced rather than heightened needs to achieve (Birney, Burdick, & Teevan, 1969).

Many kinds of personal catastrophes send people into periods of grief or depressed despair following the initial surge of protest: losses of personal relationships, roles, reputations, or "significant objects" (Averill, 1968), becoming blind (Fitzgerald, 1970) or suffering other disabilities that interfere with their pursuit or enjoyment of valued incentives (Orbach & Sutherland, 1954; Talbott, 1970), and separation from parents (Bowlby, 1960; McKinney et al., 1971). People are more likely to be hospitalized for depression when their communities offer few economic opportunities (Linsky, 1969). Depressed patients experience a much higher proportion of upsetting events within a short time before their admission than a matched group of nonclinical subjects⁵ (Paykel, Myers, Dienelt, Klerman, Lindenthal, & Pepper, 1969), and one can almost always find precipitating events in the weeks before the onset of clinical depression, usually multiple events clustered over a relatively short time period (Leff, Roatch, & Bunney, 1970; Grinker, Miller, Sabshin, Nunn, & Nunnally, 1961).

Precipitating events virtually always interfere with the patient's ability to strive for important positive incentives or with his ability to avoid threatening consequences. The most common precipitating events reported in the studies just mentioned are changes in sexual and marital relationships,

⁵ Brown and Birley (1968) have also found a strong association between life crises and the subsequent onset of schizophrenia. Since life crises can have a variety of effects, however, this does not invalidate the present argument that one prominent effect is depression.

moves, job changes, losses of loved ones, illness, and the like. When Izard (1972) asked 332 college students to describe a situation in which they were depressed, his subjects most often mentioned academic failure or the prospect thereof (22%), disrupted heterosexual relationships (21%), loneliness (15%), "loss or failure in non-academic competition" (6%), "death or illness of loved one or friend" (5%), and "difficulties with parents" (3%). The remaining 28% reported situations that combined several of these categories (10%) or reported miscellaneous other, often rather similar situations (18%).

*Depression*⁶

Whether a person's depression is mild, agitated, or retarded, he reduces or stops his instrumental activity while feeling helpless and hopeless (Bibring, 1953; Grinker et al., 1961; Wessman & Ricks, 1966; and, indeed, most clinical commentators on depression). Apathy toward normal incentives is one of the most characteristic accompaniments of depression (Beck, 1967) and grief (Marris, 1958). Depressed patients direct and receive fewer social acts and time them less skillfully (Libet & Lewinsohn, 1973; Lewinsohn, 1974). It is unclear from these data which variable, social skill or depression, is causal, but since normal subjects are less attracted to an "anonymous stranger" after having been

shown a depressing film (Gouaux, 1971) and are less attracted to excitement during depressed moods (Ludwig, 1970), and in view of the general drop in incentive values commonly observed in clinical depression (Beck, 1967), it seems likely that depression impairs social functioning by attenuating its motivational base.

Recovery

There is apparently not much evidence as to what terminates depressions other than antidepressant medications and electroconvulsive shock. Recovery usually does occur, especially from reactive depressions, even without specific therapy, but the natural precipitators of recovery are still unknown. At some point in the cycle, successes of some kind most likely play a role. Despite the fact that apathy (including loss of appetite for food and sex) is one of the pervasive clinical characteristics of depression, in at least one study depressed patients were found unusually responsive to success in minor tasks (Beck, 1967; Loeb, Beck, & Diggory, 1971), and therapy based on improving patients' social skills is claimed to have a positive effect (Lewinsohn, 1974). There is no hard evidence with which to determine whether people are especially responsive to successes throughout their depressions or whether they become increasingly responsive as their depression takes its course, a view favored by Jacobson (1971).

It would be interesting to know more about what happens to the relative value of the lost incentive during the depression phase. It seems clear enough that the value of other incentives remains low, as though there were a carry-over or perhaps even a deepening of their devaluation from the invigoration and aggression phases. However, there is clinical evidence to suppose that the value of the lost incentive remains high at least during the early part of bereavement. Parkes (1972) points out the continued searchlike behavior and affect of the bereaved. It may be that depression involves a gradual decline in the affective value of the lost incentive and a gradual shift in adaptation level to permit a recovery of other values.

⁶ The foregoing and following passages clearly tread on important diagnostic issues concerning the existence of true endogenous clinical depression (that is, without external precipitators) and concerning whether endogenous and reactive depression constitute one, two, or more dimensions. However, these issues are complex, the evidence bearing on them is inconclusive, and the issue of dimensionality is not central to the argument presented in this article. Discussion of them is therefore omitted. The argument proceeds with depression defined as an affective state or variable rather than as a diagnostic entity (although it might be inferred from a diagnostic classification). Since depression is presumably mediated by a complex system of neural and humoral components, and since a variety of organic factors can disorder this system in the absence of precipitating factors, the present formulation does not exclude possible etiologies of depression other than that under consideration here.

AN INCENTIVE-DISENGAGEMENT CYCLE

The evidence presented in the previous section demonstrates that loss of incentives leads to several specifiable effects: invigoration, primitivization of response, aggression, and depression, from which there is usually an eventual recovery. However, the data concerning the various effects come from a variety of sources: primarily experimental evidence for invigoration and aggression and primarily clinical evidence for depression. What evidence is there that these various effects of incentive loss constitute a coherent, orderly *cycle*?

Evidence for a Cycle

A few of the naturalistic and clinical studies already cited, which followed subjects longitudinally or tried to reconstruct the sequence of events retrospectively, strongly suggest the cyclical nature of the effects. Work with separation of mothers from infant monkeys (McKinney et al., 1971) and from human infants (Bowlby, 1973), and studies of bereavement (Marris, 1958; Parkes, 1972) and of blindness (Fitzgerald, 1970) all report behavior that seems to correspond to an initial stage of invigoration (activity or disbelief) followed by a stage of aggression (protest or blame) which shades into a period of depression (grief or apathy). Eventually, if the person's life situation offers some alternative incentives that seem accessible, the depression lifts and the person readjusts. The similarity of these sequences in samples of quite different ages, in different life situations, and even in different primate species suggests that the sequence is rooted in the structure of mammalian (or at least primate) motivation.

The regularity of the findings suggested that it would be worthwhile to look for such a sequence in another species with another kind of incentive (Klinger, Barta, & Kemble, 1974). Building on a finding that open-field activity of rats is higher just after unrewarded runway trials than after rewarded ones (Gallup & Altomari, 1969), the procedure measured open-field activity of hooded rats once after a runway acquisition trial and following four evenly spaced

runway trials during extinction. Four groups of rats varied as to which particular set of runway extinction trials were followed by open-field trials so that every runway extinction trial was followed by an open-field trial with one of the four groups. A fifth (control) group remained on a continuous reward schedule throughout. The critical tests of significance were the cubic component of the open-field activity trend for the four experimental groups ($p < .001$), the cubic component of the Groups $\times$ Trials interaction comparing the four combined experimental groups with the control group ($p < .025$), and the difference of the experimental groups' fourth (depressed) activity trial means from their acquisition and recovery means and from the controls' mean ($p < .05$ in each case). Clearly, all phases of the cycle except aggression are reflected in these data (see Figure 1).

The results support the theory that the incentive-disengagement sequence constitutes an orderly cycle. Thus, the phenomena of extinction, frustration, invigoration, depression, grief, and at least certain forms of aggression are brought together under a single theoretical roof within the larger framework of a particular incentive theory of motivation.

Further considerations. The fact that a certain sequence of events unfolds lawfully within certain conditions, while theoretically important, does not by itself demonstrate that the phases of the sequence are governed by a single cohesive set of mechanisms. Existing evidence is insufficient to establish the unitariness or interrelatedness of the mechanisms that might subserve such a sequence. There are, however, some indications that implicate norepinephrine and the amygdala. The evidence concerning the role of norepinephrine and the cycle is still largely speculative: Whereas aggressive states are associated with increased norepinephrine (Kety, 1970), there are indications that depression is associated with norepinephrine depletion (Schildkraut, 1965).

The evidence implicating the amygdala in incentive-related (and cycle-related) functions is considerably stronger. Ablation of the amygdaloid complex:

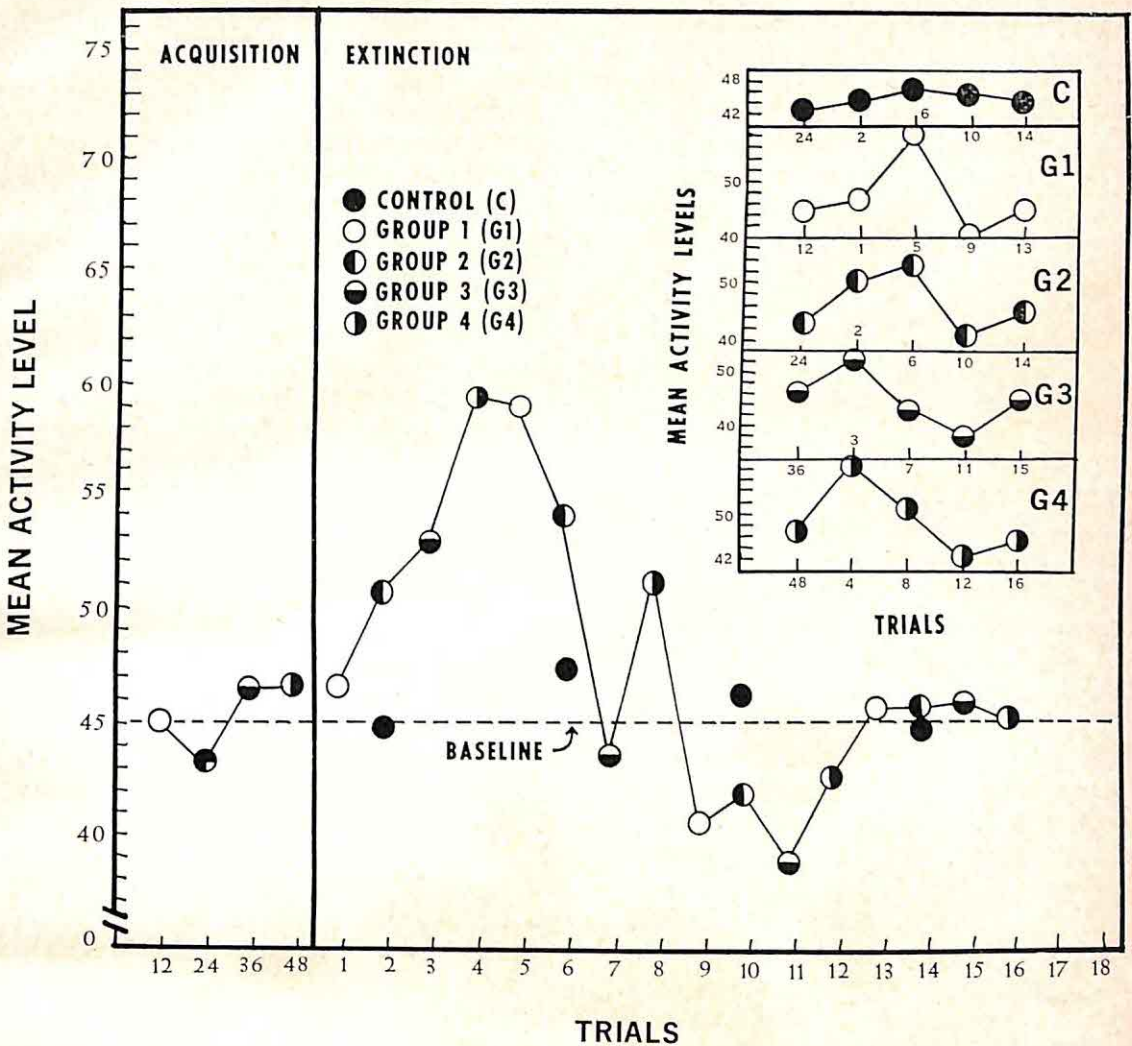


FIGURE 1. Activity levels (number of 10-cm squares entered per 60 sec) as a function of preceding runway trials. (The large curve is a composite of data from four experimental groups whose separate curves are presented in the upper right. The unconnected black dots represent the activity levels of Control Group 5, also presented in the extreme upper right. The second point from the left in the large curve represents identical activity means for Group 2 and Group 5 (control). From "Cyclic Activity Changes During Extinction in Rats: A Potential Model of Depression," by E. Klinger, E. D. Kemble, and S. G. Barta, *Animal Learning and Behavior*, 1974, 2, 313-316. Copyright 1974 by the Psychonomic Society. Reprinted by permission.)

produces profound "taming" of wild animals (e.g., Schreiner & Kling, 1953, 1956), deficiencies in a variety of aversively motivated learning tasks (e.g., Ursin, 1965, Horvath, 1963), and reduced responsivity to various positive (e.g., Kemble & Schwartzbaum, 1969, Kemble & Beckman, 1970) and negative (Kemble & Beckman, 1969) incentives. Finally electrical stimulation within the amygdala results in well-defined emotional behavior such as attack or fleeing (e.g., Ursin & Kaada, 1960). (Kemble, 1973)⁷

Mark and Ervin (1970) have reported instances in which seizure-produced rages

gave way to episodes of depression. Other clinical evidence implicates the amygdala in many such seizures. Finally, amygdectomy in rats abolishes the Amsel frustration effect (Henke & Maxwell, 1973) and all phases of the incentive-disengagement cycle

⁷ Unpublished communication (November 23, 1973) from Ernest D. Kemble, to whom I am deeply indebted for continually stimulating discussions on the neuropsychological aspects of incentive-related behavior.

as measured by open-field activity levels following runway extinction trials (Barta, Kemble, & Klinger, in press). The amygdala, of course, is a complex entity, and it participates in systems that extend anatomically well beyond it. Whatever its role may turn out to be, a distinguishing feature of the present theoretical approach is to emphasize the specific interconnectedness of the processes involved in commitment to and disengagement from incentives, and hence to prescribe a search for interconnected mechanisms.

Duration and Intensity

Incentive-disengagement cycles would seem to vary greatly in duration and intensity, ranging from minor disappointments whose effects dissipate within hours to major depressions that last for months. One might well wonder whether such quantitatively different phenomena should be regarded as manifestations of similar processes, but the similarity in their antecedents and course suggests that they should. There is, unfortunately, very little evidence concerning what determines the length and intensity of such cycles. Presumably, duration and intensity depend on individual differences and on the nature and timing of the precipitating loss. General observation suggests that greater invigoration, aggression, and depression follow the loss of more important incentives. All phases also presumably depend on cognitive factors such as the person's belief in the reality and finality of a loss, his estimate of how costly and effective aggression might be, and his assessment of the substitutability of alternative incentives. For instance, infants seem more distressed at separation from someone to whom they have formed a strong attachment, and their distress is less severe if they are quickly transferred to an effective mother surrogate (Bowlby, 1973).

Severity of a loss probably depends on the extent to which the lost incentive was capable of eliciting positive affect and on the proportion of the organism's response repertory that was shaped specifically to obtain and enjoy the incentive (Lewis, 1961; Parkes, 1972). Thus, a close personal re-

lationship or a gratifying occupation would be among the most distressing kinds of incentives to lose. The duration of the depression phase of the cycle may depend additionally on the number and prominence of the cues remaining in the organism's environment that were associated with the incentive. Such cues reduce the frustration effect (Staddon, 1970, 1972) and lengthen extinction (e.g., Melching, 1954). In bereavement, such reminders are often especially painful (Marris, 1958). Perhaps, then, they generally retard the progress of disengagement. Despite these indications, there is clearly a need for advances in both theory and research to establish the parameters of the cycle.

Adaptive Functions

Adaptive functions for individual survival. The most general characteristics of the incentive-disengagement cycle seem deducible from a priori considerations of an organism's requirements for survival. First, an organism would undoubtedly benefit from a tendency to pursue vital incentives tenaciously. A number of attributes serve this function: The pleasant states that accompany "primary reinforcement" appear to form a part of a system for establishing the psychological value of certain essential incentives; an organism's responsivity toward concern-related stimuli provides a basis for reminding it of its incentives; and, in general, the mechanisms underlying persistent states of current concern keep the organism as much as possible on target toward its goals. Moreover, if one method for attaining an incentive is blocked, the organism would clearly benefit from a tendency to attempt other methods and, if its repertory of efforts fails, to try again with greater physical and psychological force: the invigoration and aggression phases. On the other hand, such an organism would totally exhaust itself pursuing its first blocked goal if it did not incorporate a mechanism for annulling its commitment to an inaccessible incentive, thus ending that particular current concern. If, however, rest or the end of hard striving carries positive incentive value, it would be important that the disengagement process be

affectively aversive in order to prevent failure from taking on a net positive incentive value, which would surely be a dangerous state of affairs. Depression, of course, embodies the properties of such a mechanism. Within this framework, then, some degree of depression is a normal, adaptive part of the process of disengagement from incentives. Since perpetual depression would leave the organism incapacitated, there is obvious need for a recovery phase.

Adaptive functions for population survival.

In some important respects, depression is individually maladaptive. Not only are depressed individuals often hypochondriacal (Jacobs, Fogelson, & Charles, 1968), the bereaved also often experience actual deteriorating health and higher rates of morbidity (Marris, 1958) and mortality (Rees & Lutkins, 1967; Young, Benjamin, & Wallis, 1963). Life changes that require major social readjustments increase the risk of subsequent serious illnesses, especially of chronic illness (Holmes & Masuda, 1973). Most of the seven most heavily weighted events in Holmes and Masuda's index of life changes are known to produce a high incidence of depression, and most involve losses of incentives or threats of loss: death of spouse, divorce, marital separation, jail term, death of close family member, personal injury or illness, and marriage.

In the face of these grim statistics, how can the depression phase of the cycle be adaptive? Viewed from a population standpoint, there are several advantages to the demise of depressed individuals. First, severe depression is a consequence of losing a major incentive, which in a natural setting is likely to involve loss of a mate, progeny, or territory. Insofar as the loss is related to relative incompetency on the part of the individual, there is a eugenic advantage in eliminating the incompetents. Insofar as the individual has been prevented from reproducing while still laying claim to a share of the food supply, there is a reproductive advantage in his demise. Thus, a characteristic disadvantageous to certain individuals may benefit the species and may therefore be selected for in the course of evolution.

Price (1972) has suggested two other in-

triguing arguments for the biological utility of depression. First, if incentive loss can be attributed not to incompetence but to an inhospitable environment, it would be useful for an individual to respond as he might to other instances of environmental inhospitality by minimizing the metabolic cost of waiting for easier times. Price draws a parallel between "stay-put" depressions and hibernation, based in part on several behavioral similarities (including, of course, torpidity) and a few biochemical indications.

The second argument involves the observation that a number of primate species have adaptively ritualized competition for dominance, reducing potentially lethal battles to symbolic displays. However, if the loser of such a symbolic contest

has merely received signals, he is physically fit and might choose to ignore the ritual contest and engage in destructive fighting, or he might claim that the ritual was null and void in some way and insist on repeating it. This sort of behavior would counteract all the gains to be derived from ritualized fighting; in fact, the ritualization of fighting requires that the loser should yield. . . . The simplest and most effective way that he can be induced to yield is by being incapacitated. It does not matter whether he is physically or psychologically incapacitated, except that in the latter case recovery is swifter and more complete. (Price, 1972, p. 135)

Research by Gartlan & Brain (1968) and by Price (1972) has shown that the vervet monkey undergoes marked physiological and behavioral changes following a loss in a competition for social dominance, including depressed-appearing behavior and increased mortality.

Depression could, then, serve a number of adaptive functions for individuals and for species, and it may have evolved as a compromise mechanism serving several of them.

FURTHER IMPLICATIONS FOR PSYCHOLOGICAL THEORY AND APPLICATION

The concept of an incentive-disengagement cycle and related concepts articulate with a wide network of theory and practice. Their implications, therefore, extend in a number of different directions. The implications for motivational theory have already been suggested above. The present section

summarizes implications for theory and treatment of depression, for the psychology of drug use, and for other social problem areas.

Implications for a Theory of Depression

Clinical status. The formulation presented above suggests that in most cases depression is a normal, adaptive, nonpathological process which, despite its being a nuisance, need cause little concern for the psychological viability of the depressed individual. If one were to draw a physical analogy, it is more akin to a painful bruise than to a broken leg or to varicose veins. In severe cases depression, like a bruise, might call for clinical support, and it also carries an increased risk of precipitating a more serious pathology. Nevertheless, the depressive process as such can be seen as essentially part of recuperating from a loss.

This is not to deny the reality of disabling depressions. There are, no doubt, learned dispositions or constitutional factors that increase the severity of the depressive phase of the cycle or impair the person's ability to recover. Depressions that occur in conjunction with various other processes such as endocrine dysfunctions and schizophrenia require special treatment approaches. Nevertheless, the etiological view presented here suggests particular treatment strategies for "normal" depression, involving educational and supportive help to provide the person with a time perspective, to assist in an analysis of possibly maladaptive goal commitments, and to reduce anxiety, shame, etc., regarding the depression itself.

The present view raises a question as to whether a depressive phase not only accompanies disengagement from incentives but also is necessary for successful disengagement from incentives to occur. It suggests the importance of discovering to what extent various pharmacological and electroconvulsive therapies interfere with the patient's disengagement from incentives. Remedies that interfere with disengagement may have either desirable or undesirable effects, such as sending back into contention someone who is too easily discouraged or reviving a commitment to a goal that were better aban-

doned. In either case, a physical therapy unaccompanied by other help may forestall a person's examining the nature of his commitments, which may be necessary for lasting therapeutic effect.

Explanation of depressive symptoms. The formulation presented here provides direct explanations for at least four prominent accompaniments of depression: hopelessness, lowered self-esteem, links to aggression and hostility, and hypochondria. Since depression is viewed specifically here as disengagement from an incentive that seems hopelessly out of reach, the affect of hopelessness and the cognitive belief in one's helplessness can be seen as a generalization from the specific loss. If the value of the lost incentive becomes exaggerated as suggested above, the person's loss would appear to him more extensive than the original importance of the incentive warrants, especially if his sense of self-worth was predicated on effectiveness in controlling incentives.

The proposed model makes clear that both aggression and depression are products of frustration, but neither is viewed as a transformation of the other.

Finally, the concept of the cycle together with the theory of motivational effects on ideation explains the frequent finding of hypochondriacal symptoms in conjunction with depression (e.g., Jacobs et al., 1968). With the marked weakening of current concerns about external incentives, one would expect a relative increase in attention to interoceptive stimuli, which would then elicit a larger share of fantasy segments concerning them. Along with the preoccupation with bodily sensations, it is easy to see how the anxiety that often accompanies depression (Claghorn, 1970; Grinker et al., 1961) and the secondary gains that accompany the sick role would combine to produce a high incidence of hypochondria.

Mood stability and instability. Most people go about their lives with few major mood shifts, their affect varying within a tolerable range from the "blues" to mild elation, despite occasional disappointments. Clinically depressed patients are often reported to have experienced a cluster of disturbing events in

preceding months (e.g., Leff et al., 1970). The concept of an incentive-disengagement cycle can accommodate these observations with the assumption that a person can experience a separate cycle for each incentive to which he is committed and that the effects of the different cycles can summate at each moment. If, therefore, a person who has suffered a serious reverse has many other irons in the fire and perhaps something going well for him, the model would predict a dampening of elation about the things going well rather than an experienced depression. On the other hand, if the depressive phases of several cycles were to coincide in time, the model would predict some degree of depression even if none of the incentives is, by itself, of paramount importance. The model thus provides a framework for understanding fluctuations in mood.

Individual differences. The effects of the incentive-disengagement cycle probably depend on individual differences in constitution (Rosenthal, 1970), cognition, and other consequences of learning histories. The concept of the incentive-disengagement cycle provides a framework for analyzing the dimensions and sources of individual differences in susceptibility to depression. The five "phases" of the cycle each suggests different sources. For instance, the invigoration phase suggests that the onset of the aggressive phase may be hastened by possession of a limited response repertory or the phase may be truncated by interpretations that lead to early hopelessness such as certain kinds of causal attributions. The nature of the aggression phase suggests that histories of reward or punishment for aggressive behavior affect its duration since they would affect the amount of unrewarded effort necessary to crush hope. The two phases jointly suggest a search for factors that make people differentially extinguishable. For instance, Amsel (1972) has recently reported research on extinguishability and persistence as a function of partial reinforcement training in rats. Dweck and Repucci (1973) have found that children who attribute success or failure to their own efforts are less likely to succumb to "learned

helplessness." Under extinction conditions, subjects who attribute their failure to low ability or task difficulty become much more pessimistic about their chances of success than subjects who attribute their failure to lack of effort or bad luck (Weiner, Heckhausen, Meyer, & Cook, 1972). The depressive downswing phase suggests a search for constitutional factors and also for factors involving number and variety of a person's simultaneous commitments to incentives, since it seems likely that a downswing will be less severe if the individual has many other promising commitments. Finally, the recovery phase suggests a search for factors that help the depressed individual find substitute incentives and gratifications.

Inner experience during bereavement and depression. "The most characteristic feature of grief," writes Parkes (1972), "is not prolonged depression but acute and episodic 'pangs'. . . . At such a time the lost person is strongly missed and the survivor sobs or cries aloud for him" (p. 39). Furthermore, bereaved people frequently find themselves thinking thoughts of a kind that had become habitual (such as whether a mate will like something the bereaved person is doing) but which are now inappropriate (Marris, 1958). When the theory of depression presented here is combined with the theory of inner experience detailed elsewhere (Klinger, 1971, Note 1, Note 2), the larger theoretical structure can account for these phenomena. First, bereavement falls within a period of disengagement during which the person still retains a current concern about the lost individual. Describing both clinical and ethological evidence, Parkes (1972) concludes that "[pinning] is the subjective and emotional component of the urge to search for a lost object" (p. 40). Second, however, segments of thought about a concern, along with their associated affect, seem to be induced when the person encounters cues (both sensory and ideational) related to those concerns. Third, the segments that then unfold are integrated response sequences established through previous overlearning. Thus, within this framework, segments that occur as if the lost person were still alive can be under-

stood as the perservation of integrated sequences induced by a still-alive concern and customary cues. Similarly, pangs of grief can be viewed as emotionally laden segments induced in the same way as other segments of thought.

Since depression carries with it indifference to most incentives, it involves a weakening of most current concerns. If current concerns govern responsiveness to cues and govern the content of fantasy, the model predicts that the severely depressed individual will be relatively unresponsive to most stimuli since most stimuli are irrelevant to him, even though sensory imagery may make up a large share of his conscious experience. In addition to a drastically reduced level of operant activity, the depressed person's weaker affective responsiveness to most external stimuli should lead to his engaging in more repetitive, stale, drab fantasy, much of it perhaps cued off by somatic sensations and bodily concerns. The model thus provides some specific testable hypotheses regarding depressive inner experience. Since in accordance with the larger model these factors similarly influence operant thought, the model also can generate specific predictions regarding the impairment of the depressed person's ability to cope with the exigencies of his life situation.

The theory of depression as learned helplessness. Recently, Seligman (1974) proposed a theory of reactive depression as "learned helplessness." The theory grew out of observations that animals subjected to inescapable electric shocks in one situation (a Pavlovian hammock) were slower to initiate escape responses in a second situation (a shuttlebox) and slower to learn escape responses after beginning to emit them (Overmeir & Seligman, 1967; Seligman & Groves, 1970; Seligman & Maier, 1967). Other studies have similarly found that suffering inescapable discomfort retards escape learning in humans (Hiroto, 1971; Racinskis, 1971), although one study (Thornton & Jacobs, 1971) found no significant differences between an inescapably shocked group and a nonshocked control group. Seligman explains the subjects' impaired escape learning as arising from learning that the aver-

sive stimulation is noncontingent on their responses—that they are helpless to prevent discomfort. Presumably, learned helplessness is, in this view, the critical event that triggers reactive depression. Like depressed individuals, organisms subjected to inescapable stress are passive and apparently pessimistic. Both have experienced uncontrollable stress. There is some evidence that both conditions are accompanied by norepinephrine depletion. Primarily on the basis of these classes of evidence, Seligman proposes "learned helplessness" as a model of reactive depression.

Seligman's position and the present theory agree that both extinction and inescapable stress are depressing. Seligman's evidence rests primarily on the effects of inescapable stress, but his theory considers the operative factor to be lack of control over reinforcers, which includes extinction of appetitive behavior. The theory of the incentive-disengagement cycle rests on evidence concerning the effects of appetitive extinction, failure, and incentive losses, but these categories logically include losses of comfort and failures to ensure safety. Seligman stresses that the sense of control over events is itself valued. The theory of current concerns stresses the organizing effect of contingent incentives on behavior and on the content of thought and affect. Thus, both positions agree concerning the *importance* of an incentive's controllability, although the reasons for regarding it as important seem to differ.

Apart from these similarities, there are at least three areas of disagreement. First, is the passivity of a depressed organism due to its having learned it is helpless or is it an unlearned accompaniment of a more comprehensive state? Second, is that state precipitated by helplessness as such or by the loss of incentives? Third, can teaching a person confidence in his powers to control events ("mastery training") reduce the depression he experiences following losses? These are important issues that bear some discussion.

Concerning the *first of these issues*, there are at least two kinds of observations that seem to pose problems for the view that the passivity of depression is learned directly.

First, depression afflicts many people who, despite their loss of some significant incentive, had been succeeding in their pursuit of one or more other significant incentives. Thus, the person who loses a mate but remains successful in his or her work, or who loses a job while remaining on good terms with loved ones, may become depressed despite evidence of his or ability to control various incentives. Why should his sense of helplessness in one restricted life sector generalize to so many other sectors?

Even more difficult to accommodate within the framework of learned helplessness, the recovery phase of the incentive-disengagement cycle has been observed in rats that were still being subjected to continuing extinction trials (Klinger et al., 1974; see Figure 1). If their depressed levels of open-field activity are consequences of learned helplessness, why did the depression dissipate while the lesson of their helplessness continued—that is, while their runaway speeds continued to fall?

The *second major issue* concerns the precipitating circumstances of depression: helplessness or the loss of incentives. One experiment with human subjects (Roth & Bootzin, 1974) produced helplessness by providing random feedback in a concept learning task without using aversive stimulation. These "helpless" subjects later were *more* likely than controls to take steps against equipment malfunctions.

The learned-helplessness position assumes that helplessness (i.e., noncontingency of outcomes on behavior) is inherently aversive. However, such a proposition appears hard to defend. Whether noncontingency is aversive clearly depends on what is involved. Pigeons and rats are willing to work for periods of noncontingent reward (Neuringer, 1973), and among humans commercial gambling finds many devotees.

A *third major issue* concerns the effectiveness of a preventive approach suggested by the learned-helplessness position, the notion that "mastery training" in childhood can "immunize" people against depression later. Animals with previous successful experiences at escaping stress, who are then subjected to inescapable stress, later relearn

escape, when that becomes possible again, more easily than animals whose inescapable stress constitutes their first experimental experience. Experience with partial reinforcement schedules can also increase rats' persistence during extinction (Amsel, 1972). However, there is, so far, no indication that people can be immunized against depression. (Why does success not immunize against success depressions?) The incentive-disengagement theory does agree that early mastery training may reduce the frequency of depression by teaching people to persist longer during the invigoration and aggression phases of the cycle, thereby preventing premature discouragement and failures by default. However, nothing in the incentive-disengagement position suggests that mastery training would reduce the severity of depression when losses occur. Therefore, this viewpoint suggests that it may be desirable to educate people concerning the course, inevitability, and naturalness of depressive reactions, thereby perhaps alleviating some of the secondary consequences of depression, such as anxiety and self-doubts aroused by the fact of depression having occurred.

If the incentive-disengagement theory of depression is correct, how can we explain the results of inescapable stress? It may be that these reflect the operation of two factors. First, subjects in these experiments may indeed learn the uselessness of responding, just as the learned-helplessness theory suggests. This, then, would account for the subject's subsequent passivity in the face of opportunities to escape. Second, however, the removal of such incentives as comfort and safety may also produce a state of depression, which may then account for some of the more general effects, similar to those of other incentive loss. In other words, the two-factor explanation suggests that incentive loss produces depression. Those situations in which the organism is unable to escape stress additionally produce specific learned helplessness.

Implications for Drug Use

The incentive-disengagement cycle consists of a succession of feeling states, most

of them unpleasant. Psychoactive drugs provide ways to alter feeling states. Although some effects of drugs seem to be producible by any of a number of substances—for instance, drug-produced “highs” seem to have some common elements—different varieties of drugs exert different specific effects as well. For instance, marijuana enhances the incentive value of almost all sensory experiences (McKenzie, Note 3; Tart, 1971). Alcohol in at least moderate dosages confers an increased sense of optimism and “power” (McClelland, Davis, Kalin, & Wanner, 1972), as do the amphetamines, which additionally reduce fatigue, depression, and boredom (Cole, 1967). Marlatt (Note 4) reports that “almost a third of all relapses [of alcoholic patients] occurred when the patient was frustrated or angry;” and he reports experimental findings that heavy drinkers drink heavily after they have been insulted with no opportunity to retaliate, but they drink little if they had an opportunity to retaliate. Scattered interview evidence suggests that among long-term users heroin has a normalizing effect on mood, anesthetizing the user to the effects of living with few major positive incentives and numerous negative ones (Cortina, 1970).

The possibility therefore arises that the specific effects of various drug classes may be differentially rewarding in different phases of the incentive-disengagement cycle. In that case, the concept of the cycle may help to account for differential patterns of use. For instance, one might suppose that marijuana would appeal especially to people who seek to enhance incentive values, perhaps because they are deprived of appealing incentives, because they have become too adapted to the incentives available to them, or for recreational purposes. Similarly, alcohol may be especially preferred during the latter portions of the invigoration or aggression stages, perhaps especially when the individual fears failure or impotence due to his own personal limitations. Heroin may become a drug of choice in desperate situations barren of hope for positive incentives and rich in negative ones; and so on.

This is not, of course, to lose sight of the multiple factors that determine drug use, such as the general pleasure of a “high,” cultural factors governing use, availability, costs, ritual, and beliefs in self-exploratory or other cognitive and experiential benefits. Nevertheless, the conception of the incentive-disengagement cycle suggests a powerful array of factors for the analysis of differential drug preferences and use patterns.

Other Implications

The conceptualization described above can be applied to a number of other problems. For instance, changes in age status often force people to give up satisfying activities and enlivening aspirations. Changes in the value of old incentives, such as jobs and marriage partners, may turn an incentive-rich life situation into a barren one. Although the consequences of these relatively gradual changes in incentive value are not discussed above, they may have similar effects as losses, including depression (Costello, 1972). Chronically barren life situations probably produce rather different outcomes, but here, too, in view of the organizational properties of current concerns, the present theory can contribute predictions.

The theoretical position advanced here also carries a number of implications for education, psychotherapy, and eupsychian management. It suggests that one of the functions of education for mental health should be to teach the nature and ramifications of commitments to goals, including both the virtues and the risks of aspiring to difficult goals. Such education would place the experience of depression into a comprehensible framework that is neither alarmist nor dismissing but that treats depression as an important kind of information about one's life situation from which one can deduce new plans for self-fulfillment. Psychotherapy for depression would incorporate these functions and would additionally aid patients in working out coping strategies suitable to their situation. Finally, the present theory suggests managerial techniques to improve employees' functioning and job satisfaction.

CONCLUSIONS

There is compelling evidence that behavior is organized around the enjoyment and pursuit of incentives. Between the time an organism commits itself to striving for an incentive and the time it consummates or abandons the incentive it is in a special state with respect to the incentive, a state called here a "current concern." During states of current concern, organisms are more responsive to concern-related stimuli than they would otherwise be to those stimuli, and they respond with elements of the concern-related responses in their repertory. Therefore, current concerns potentiate concern-related content in fantasy. Loss or impoverishment of a person's incentives can therefore further be predicted to reduce activity level, flatten or depress affect, reduce the variety and interest of fantasy, and promote stimulus-bound and especially somatic fantasy.

When organisms are blocked in their pursuit of an incentive, they go through an orderly sequence of changes called here the *incentive-disengagement cycle*, whose phases (which are not discrete) include invigoration of behavior, aggression, downswing into depression, depression, and recovery. The content of the cycle in a particular instance depends in predictable ways on the particular circumstances, and the severity of affect depends on the importance of the incentive. The cycle corresponds to the course of extinction. Depression is ordinarily therefore a normal, adaptive part of disengaging oneself from an incentive.

The concept of the incentive-disengagement cycle suggests places to look for individual differences in susceptibility to depression and can account for depressive symptoms such as hopelessness, hostility, lowered self-esteem, and hypochondria. The particular kinds of unpleasant fantasy and affect experienced during the incentive-disengagement cycle may mesh with the specific effects of psychoactive drugs to help explain the use of marijuana, alcohol, amphetamines, narcotics, and probably other "mind-altering" substances.

The effective value of an incentive changes

with changes in the expectation that it can be attained, with habituation, changes in adaptation level, and experience with the incentive class. These factors may operate to impoverish a person's incentive world and to affect the course of careers, marriage, and aging.

The incentive-disengagement theory of depression carries specific implications for psychotherapy, education, and eupsychian management.

REFERENCE NOTES

1. Klinger, E. *Meaning and void: Inner experience and the incentive world*. Manuscript in preparation, 1975.
2. Klinger, E. *The nature of fantasy and its clinical uses*. Manuscript submitted for publication, 1975.
3. McKenzie, J. D. *Trends in marijuana use among undergraduate students at the University of Maryland* (Counseling Center Research Report No. 3-70). College Park: University of Maryland Counseling Center, 1970.
4. Marlatt, G. A. Determinants of alcohol consumption in a laboratory taste-rating task: Implications for controlled drinking. In Cyril Franks (Chair), *Can alcoholics learn to control their drinking?* Symposium presented at the 81st Annual Meeting of the American Psychological Association, Montreal, Canada, 1973.

REFERENCES

- Adamson, R. Contrast effects and reinforcement. In M. H. Appley (Ed.), *Adaptation level theory*. New York: Academic, 1971.
- Amsel, A. The role of frustrative nonreward in noncontinuous reward situations. *Psychological Bulletin*, 1958, 55, 102-119.
- Amsel, A. Frustrative nonreward in partial reinforcement and discrimination learning: Some recent history and a theoretical extension. *Psychological Review*, 1962, 69, 306-328.
- Amsel, A. Behavioral habituation, counterconditioning, and a general theory of persistence. In A. H. Black & W. F. Prokasy (Eds.), *Classical conditioning II*. New York: Appleton-Century-Crofts, 1972.
- Appley, M. H. Derived motives. *Annual Review of Psychology*, 1970, 21, 485-511.
- Appley, M. H. (Ed.). *Adaptation-level theory*. New York: Academic, 1971.
- Atkinson, J. W. *An introduction to motivation*. Princeton, N. J.: Van Nostrand, 1964.
- Atkinson, J. W., & Birch, D. *The dynamics of action*. New York: Wiley, 1970.
- Atkinson, J. W., & McClelland, D. C. The projective expression of needs. II. The effect of

- different intensities of the hunger drive on thematic apperception. *Journal of Experimental Psychology*, 1948, 38, 643-658.
- Atkinson, J. W., & Raynor, J. O. *Motivation and achievement*. Washington, D.C.: V. H. Winston, 1974.
- Atkinson, R. C., & Wickens, T. D. Human memory and the concept of reinforcement. In R. Glaser (Ed.), *The nature of reinforcement*. New York: Academic Press, 1971.
- Averill, J. R. Grief: Its nature and significance. *Psychological Bulletin*, 1968, 70, 721-748.
- Azrin, N. H., Hutchinson, R. R., & Hake, D. F. Extinction-induced aggression. *Journal of the Experimental Analysis of Behavior*, 1966, 9, 191-204.
- Barker, R. H., Dembo, T., & Lewin, K. Frustration and regression: An experiment with young children. *University of Iowa Studies in Child Welfare*, 1941, 18(1).
- Barta, S. G., Kemble, E. D., & Klinger, E. Abolition of cyclic activity changes following amygdaloid lesions in rats. *Bulletin of the Psychonomic Society*, 1975, in press.
- Beck, A. T. *Depression: Clinical, experimental, and theoretical aspects*. New York: Harper & Row, 1967.
- Berger, D. F. Alternative interpretations of the frustration effect. *Journal of Experimental Psychology*, 1969, 81, 475-483.
- Berkowitz, L. *Aggression: A social psychological analysis*. New York: McGraw-Hill, 1962.
- Berlyne, D. E. Arousal and reinforcement. In D. Levine (Ed.), *Nebraska Symposium on Motivation* (Vol. 15). Lincoln: University of Nebraska Press, 1967.
- Bibring, E. The mechanism of depression. In P. Greenacre (Ed.), *Affective disorders*. New York: International Universities Press, 1953.
- Bindra, D. Neuropsychological interpretation of the effects of drive and incentive-motivation on general activity and instrumental behavior. *Psychological Review*, 1968, 75, 1-22.
- Bindra, D. A motivational view of learning, performance, and behavior modification. *Psychological Review*, 1974, 81, 199-213.
- Birney, R. C., Burdick, H., & Teevan, R. C. *Fear of failure*. New York: Van Nostrand Reinhold, 1969.
- Bolles, R. C. *Theory of motivation*. New York: Harper & Row, 1967.
- Bolles, R. C. Reinforcement, expectancy, and learning. *Psychological Review*, 1972, 79, 394-409.
- Bowlby, J. Grief and mourning in infancy and early childhood. In R. S. Eissler, A. Freud, H. Hartmann, & M. Kris (Eds.), *Psychoanalytic study of the child* (Vol. 15). New York: International Universities Press, 1960.
- Bowlby, J. *Attachment and loss*. (Vol. 2). *Separation: Anxiety and danger*. New York: Basic Books, 1973.
- Brehm, J. W. *Responses to loss of freedom: A theory of psychological reactance*. Morristown, N. J.: General Learning Press, 1972.
- Brown, G. W., & Birley, J. L. T. Crises and life changes and the onset of schizophrenia. *Journal of Health and Social Behavior*, 1968, 9, 203-214.
- Bruner, J. S., & Goodman, C. C. Value and need as organizing factors in perception. *Journal of Abnormal and Social Psychology*, 1947, 42, 33-44.
- Butterfield, E. C. The interruption of tasks: Methodological, factual, and theoretical issues. *Psychological Bulletin*, 1964, 62, 309-322.
- Claghorn, J. The anxiety-depression syndrome. *Psychosomatics*, 1970, 11, 438-441.
- Cofer, C. N., & Appley, M. H. *Motivation: Theory and research*. New York: Wiley, 1964.
- Cole, S. O. Experimental effects of amphetamines: A review. *Psychological Bulletin*, 1967, 68, 81-90.
- Corteen, R. S., & Wood, B. Autonomic responses to shock-associated words in an unattended channel. *Journal of Experimental Psychology*, 1972, 94, 308-313.
- Cortina, F. M. *Stroke a slain warrior*. New York: Columbia University Press, 1970.
- Costello, C. G. Depression: Loss of reinforcers or loss of reinforcer effectiveness? *Behavior Therapy*, 1972, 3, 240-247.
- Dembo, T. Der Ärger als dynamisches Problem. *Psychologische Forschung*, 1931, 15, 1-144.
- DiLollo, V., Davidson, R. C., Hammond, G. R., & Donovan, R. J. The relative effects of non-reward and response thwarting on the frustration effect. *Psychonomic Science*, 1968, 10, 31-32.
- Dollard, J., et al., *Frustration and aggression*. New Haven, Conn.: Yale University Press, 1939.
- Dweck, C. S., & Repucci, N. D. Learned helplessness and reinforcement responsibility in children. *Journal of Personality and Social Psychology*, 1973, 25, 109-116.
- Edwards, W. The theory of decision making. *Psychological Bulletin*, 1954, 51, 380-417.
- Ellner, M. R. *The effects of experimentally induced frustration on depressed and nondepressed college students*. Unpublished doctoral dissertation, City University of New York, 1970.
- Erdelyi, M. H. A new look at the new look: Perceptual defense and vigilance. *Psychological Review*, 1974, 81, 1-25.
- Estes, W. K. Reinforcement in human learning. In J. T. Tapp (Ed.), *Reinforcement and behavior*. New York: Academic Press, 1969.
- Estes, W. K. Reward in human learning: Theoretical issues and strategic choice points. In R. Glaser (Ed.), *The nature of reinforcement*. New York: Academic Press, 1971.
- Feather, B. Semantic generalization in classically conditioned responses: A review. *Psychological Bulletin*, 1965, 63, 425-441.
- Festinger, L. A theoretical interpretation of shifts in level of aspiration. *Psychological Review*, 1942, 49, 235-250.

- Fitzgerald, R. C. Reactions to blindness: An exploratory study of adults with recent loss of sight. *Archives of General Psychiatry*, 1970, 22, 370-379.
- Forehand, R., & Baumeister, A. A. Rate of stereotyped body rocking of severe retardates as a function of frustration of goal-directed behavior. *Journal of Abnormal Psychology*, 1971, 78, 35-42.
- Fowler, H. Implications of sensory reinforcement. In R. Glaser (Ed.), *The nature of reinforcement*. New York: Academic Press, 1971.
- Friedman, A. S. Hostility factors and clinical improvement in depressed patients. *Archives of General Psychiatry*, 1970, 23, 524-537.
- Gallup, G. G., Jr. Aggression in rats as a function of frustrative nonreward in a straight alley. *Psychonomic Science*, 1965, 3, 99-100.
- Gallup, G. G., Jr., & Altomari, E. Activity as a postsituation measure of frustrative nonreward. *Journal of Comparative and Physiological Psychology*, 1969, 68, 382-384.
- Gartlan, J. S., & Brain, C. K. Ecology and social variability in *Cercopithecus aethiops* and *C. mitis*. In P. C. Jay (Ed.), *Primates: Studies in adaptation and variability*. New York: Holt, Rinehart & Winston, 1968.
- Gentry, W. D. Sex differences in the effects of frustration and attack on emotion and vascular processes. *Psychological Reports*, 1970, 27, 383-390.
- Gerard, H. B., Green, D., Hoyt, M., & Conolley, E. S. Influence of affect on exposure-frequency estimates. *Proceedings of the 80th Annual Convention of the American Psychological Association*, 1972, 7, 173-174.
- Gilbert, D. H. *Reward expectancy strength as related to the magnitude of frustration in children*. Unpublished doctoral dissertation, University of Southern California, 1969.
- Gilchrist, J. C., & Nesberg, L. S. Need and perceptual change in need-related objects. *Journal of Experimental Psychology*, 1952, 44, 369-376.
- Gouaux, C. Induced affective states and interpersonal attraction. *Journal of Personality and Social Psychology*, 1971, 20, 37-43.
- Grinker, R. R., Miller, J., Sabshin, M., Nunn, R., & Nunnally, J. C. *The phenomena of depressions*. New York: Harper & Row, 1961.
- Haber, R. N. Nature of the effect of set on perception. *Psychological Review*, 1966, 73, 335-351.
- Heckhausen, H. Achievement motive research: Current problems and some contributions toward a general theory of motivation. In W. J. Arnold (Ed.), *Nebraska Symposium on Motivation* (Vol. 16). Lincoln: University of Nebraska Press, 1968.
- Henke, P. G., & Maxwell, D. Lesions in the amygdala and the frustration effect. *Physiology and Behavior*, 1973, 10, 647-650.
- Hinton, B. L. Environmental frustration and creative problem solving. *Journal of Applied Psychology*, 1968, 52, 211-217.
- Hiroto, D. S. *The relationship between learned helplessness and locus of control*. Unpublished doctoral dissertation, University of Portland, 1971.
- Holmes, T. H., & Masuda, M. Life changes and illness susceptibility. In J. P. Scott & E. C. Senay (Eds.), *Separation and depression: Clinical and research aspects*. Washington, D.C.: American Association for the Advancement of Science, 1973.
- Holton, R. B. Amplitude of an instrumental response following the cessation of reward. *Child Development*, 1961, 32, 107-116.
- Horvath, F. E. Effects of basolateral amygdectomy on three types of avoidance behavior in cats. *Journal of Comparative and Physiological Psychology*, 1963, 56, 380-389.
- Hull, C. L. *A behavior system: An introduction to behavior theory concerning the individual organism*. New Haven, Conn.: Yale University Press, 1952.
- Irwin, F. W. *Intentional behavior and motivation: A cognitive theory*. New York: Lippincott, 1971.
- Izard, C. E. *Patterns of emotions: A new analysis of anxiety and depression*. New York: Academic Press, 1972.
- Jacobs, T. J., Fogelson, S., & Charles, E. Depression ratings in hypochondria. *New York State Journal of Medicine*, 1968, 68, 3119-3122.
- Jacobson, E. *Depression*. New York: International Universities Press, 1971.
- Jenkin, N. Affective processes in perception. *Psychological Bulletin*, 1957, 54, 100-127.
- Kelly, J. F., & Hake, D. F. An extinction-induced increase in an aggressive response with humans. *Journal of the Experimental Analysis of Behavior*, 1970, 14, 153-164.
- Kemble, E. D., & Beckman, G. J. Escape latencies at three levels of electric shock in rats with amygdaloid lesions. *Psychonomic Science*, 1969, 14, 205-206.
- Kemble, E. D., & Beckman, G. J. Runway performance of rats following amygdaloid lesions. *Physiology and Behavior*, 1970, 5, 45-47.
- Kemble, E. D., & Schwartzbaum, J. S. Reactivity to taste properties of solutions following amygdaloid lesions. *Physiology and Behavior*, 1969, 4, 981-985.
- Kety, S. S. Neurochemical aspects of emotional behavior. In P. Black (Ed.), *Physiological correlates of emotion*. New York: Academic, 1970.
- Klinger, E. *Structure and functions of fantasy*. New York: Wiley, 1971.
- Klinger, E. Models, context, and achievement fantasy: Parametric studies and theoretical propositions. *Journal of Personality Assessment*, 1973, 37, 25-47.
- Klinger, E., Barta, S. G., & Kemble, E. D. Cyclic activity changes during extinction in rats: A

- potential model of depression. *Animal Learning and Behavior*, 1974, 2, 313-316.
- Knott, P. D., Nunnally, J. C., & Duchnowski, A. J. Effects of frustration on primary and conditioned incentive value. *Journal of Experimental Research in Personality*, 1967, 2, 140-149.
- Lambert, W. W., Solomon, R. L., & Watson, P. D. Reinforcement and extinction as factors in size estimation. *Journal of Experimental Psychology*, 1949, 37, 637-641.
- Leff, M. J., Roatch, J. F., & Bunney, W. E., Jr. Environmental factors preceding the onset of severe depressions. *Psychiatry*, 1970, 33, 293-311.
- Levine, M. Hypothesis theory and nonlearning despite ideal S-R-reinforcement contingencies. *Psychological Review*, 1971, 78, 130-140.
- Lewin, K. Wille, Vorsatz, und Bedürfnis. *Psychologische Forschung*, 1928, 7, 330-385.
- Lewinsohn, P. M. A behavioral approach to depression. In R. J. Friedman & M. M. Katz (Eds.), *The psychology of depression: Contemporary theory and research*. New York: Winston/Wiley, 1974.
- Lewis, C. S. *A grief observed*. London: Faber & Faber, 1961.
- Libet, J. M., & Lewinsohn, P. M. Concept of social skill with special reference to the behavior of depressed persons. *Journal of Consulting and Clinical Psychology*, 1973, 40, 304-312.
- Linsky, A. S. Community structure and depressive disorders. *Social Problems*, 1969, 17, 120-131.
- Lish, J. A. *The influence of oral dependency, failure, and social exposure upon self-esteem and depression*. Unpublished doctoral dissertation, New York University, 1969.
- Loeb, A., Beck, A. T., & Diggory, J. Differential effects of success and failure on depressed and nondepressed patients. *Journal of Nervous and Mental Disease*, 1971, 152, 106-114.
- Logan, F. A. Frustration effect following correlated nonreinforcement. *Journal of Experimental Psychology*, 1968, 78, 396-400.
- Logan, F. A. *Incentive*. New Haven, Conn.: Yale University Press, 1960.
- Logan, F. A. Incentive theory, reinforcement and education. In R. Glaser (Ed.), *The nature of reinforcement*. New York: Academic Press, 1971.
- Ludwig, L. D. Intra- and interindividual relationships between elation-depression and desire for excitement. *Journal of Personality*, 1970, 32, 167-176.
- Luria, A. R., & Vinogradova, O. S. An objective investigation of the dynamics of semantic systems. *British Journal of Psychology*, 1959, 50, 89-105.
- Mandler, G. The interruption of behavior. In D. Levine (Ed.), *Nebraska Symposium on Motivation* (Vol. 12). Lincoln: University of Nebraska Press, 1964.
- Mark V. H., & Ervin, F. R. *Violence and the brain*. New York: Harper & Row, 1970.
- Marris, P. *Widows and their families*. London: Routledge & Keegan Paul, 1958.
- McClelland, D. C., & Atkinson, J. W. The projective expression of needs. I. The effect of different intensities of the hunger drive on perception. *Journal of Psychology*, 1948, 25, 205-232.
- McClelland, D. C., Atkinson, J. W., Clark, R. A., & Lowell, E. L. *The achievement motive*. New York: Appleton-Century-Crofts, 1953.
- McClelland, D. C., Davis, W. N., Kalin, R., & Wanner, E. *The drinking man: Alcohol and human motivation*. New York: Free Press, 1972.
- McKinney, W. T., Jr., Suomi, S. J., & Harlow, H. F. Depression in primates. *American Journal of Psychiatry*, 1971, 127, 1313-1320.
- Melching, W. H. The acquired reward value of an intermittently presented neutral stimulus. *Journal of Comparative and Physiological Psychology*, 1954, 47, 370-374.
- Miller, G. A., Galanter, E., & Pribram, K. H. *Plans and the structure of behavior*. New York: Holt, Rinehart & Winston, 1960.
- Mischel, W., & Masters, J. C. Effects of probability of reward attainment on responses to frustration. *Journal of Personality and Social Psychology*, 1966, 3, 390-396.
- Mitchell, T. R., & Biglan, A. Instrumentality theories: Current uses in psychology. *Psychological Bulletin*, 1971, 76, 432-454.
- Moray, N. Attention in dichotic listening: Affective cues and the influence of instructions. *Quarterly Journal of Experimental Psychology*, 1959, 11, 56-60.
- Morgan, C. T. *Physiological psychology*. New York: McGraw-Hill, 1943.
- Murray, H. A. *Explorations in personality*. New York: Oxford, 1938.
- Neisser, U. *Cognitive psychology*. New York: Appleton-Century-Crofts, 1967.
- Neuringer, A. Pigeons respond to produce periods in which rewards are independent of responding. *Journal of Experimental Analysis of Behavior*, 1973, 19, 39-54.
- Nuttin, J., & Greenwald, A. G. *Reward and punishment in human learning*. New York: Academic Press, 1968.
- Orbach, C. E., & Sutherland, A. M. Acute depressive reactions to surgical treatment for cancer. In P. H. Hoch & J. Zubin (Eds.), *Depression*. New York: Grune & Stratton, 1954.
- Oswald, I., Taylor, A. M., & Treisman, M. Discriminative responses to stimulation during human sleep. *Brain*, 1960, 83, 440-453.
- Overmier, J. B., & Seligman, M. E. P. Effects of inescapable shock upon subsequent escape and avoidance responding. *Journal of Comparative and Physiological Psychology*, 1967, 63, 28-33.
- Ovsiankina, M. Die Wiederaufnahme unterbrochener Handlungen. *Psychologische Forschung*, 1928, 2, 302-379.

- Parkes, C. M. *Bereavement: Studies of grief in adult life*. New York: International Universities Press, 1972.
- Paykel, E. S., Myers, J. K., Dienelt, M. M., Klerman, G. L., Lindenthal, J. J., & Pepper, M. P. Life events and depression. *Archives of General Psychiatry*, 1969, 21, 753-760.
- Perkins, C. C., Jr. An analysis of the concept of reinforcement. *Psychological Review*, 1968, 75, 155-172.
- Postman, L. The experimental analysis of motivational factors in perception. In J. S. Brown (Ed.), *Nebraska Symposium on Motivation* (Vol. 1). Lincoln: University of Nebraska Press, 1953.
- Price, J. S. Genetic and phylogenetic aspects of mood variation. *International Journal of Mental Health*, 1972, 1, 124-144.
- Racinkas, J. R. *Maladaptive consequences of loss or lack of control over aversive events*. Unpublished doctoral dissertation, University of Waterloo, Canada, 1971.
- Razran, C. *Mind in evolution: An East-West synthesis of learned behavior and cognition*. Boston: Houghton Mifflin, 1971.
- Rees, W. D., & Lutkins, S. G. Mortality of bereavement. *British Medical Journal*, 1967, 4, 13-16.
- Rosenthal, D. *Genetic theory and abnormal behavior*. New York: McGraw-Hill, 1970.
- Roth, S., & Bootzin, R. R. Effects of experimentally induced expectancies of external control: An investigation of learned helplessness. *Journal of Personality and Social Psychology*, 1974, 29, 253-264.
- Ryan, T. J., & Watson, P. Frustrative nonreward theory applied to children's behavior. *Psychological Bulletin*, 1968, 69, 111-125.
- Saugstad, P. The effect of food-deprivation on perception-cognition. *Psychological Bulletin*, 1966, 65, 80-90.
- Schildkraut, J. J. The catecholamine hypothesis of affective disorders: A review of supporting evidence. *American Journal of Psychiatry*, 1965, 122, 509-522.
- Schreiner, L., & Kling, A. Behavioral changes following rhinencephalic injury in cat. *Journal of Neurophysiology*, 1953, 16, 643-689.
- Schreiner, L., & Kling, A. Rhinencephalon and behavior. *American Journal of Physiology*, 1956, 184, 486-490.
- Seligman, M. E. P. Depression and learned helplessness. In R. J. Friedman & M. M. Katz (Eds.), *The psychology of depression: Contemporary theory and research*. New York: Winston/Wiley, 1974.
- Seligman, M. E. P., & Groves, D. P. Nontransient learned helplessness. *Psychonomic Science*, 1970, 19, 191-192.
- Seligman, M. E. P., & Maier, S. F. Failure to escape traumatic shock. *Journal of Experimental Psychology*, 1967, 74, 1-9.
- Sheffield, F. D., Roby, T. B., & Campbell, B. A. Drive reduction versus consummatory behavior as determinants of reinforcement. *Journal of Comparative and Physiological Psychology*, 1954, 47, 349-354.
- Spence, K. W. *Behavior theory and conditioning*. New Haven, Conn.: Yale University Press, 1956.
- Staddon, J. E. R. Temporal effects of reinforcement: A negative "frustration" effect. *Learning and Motivation*, 1970, 1, 227-247.
- Staddon, J. E. R. Reinforcement omission on temporal go-no-go schedules. *Journal of the Experimental Analysis of Behavior*, 1972, 18, 223-229.
- Talbott, J. A. Phenothiazine toxicity in pool shark. *New York State Journal of Medicine*, 1970, 70, 1671-1672.
- Tapp, J. T. Activity, reactivity, and the behavior-directing properties of stimuli. In J. T. Tapp (Ed.), *Reinforcement and behavior*. New York: Academic Press, 1969.
- Tart, C. T. Waking from sleep at a preselected time. *Journal of the American Society of Psychosomatic Dentistry and Medicine*, 1970, 17, 3-16.
- Tart, C. T. *On being stoned: A psychological study of marijuana intoxication*. Palo Alto, Calif.: Science and Behavior Books, 1971.
- Thornton, J. W., & Jacobs, P. D. Learned helplessness in human subjects. *Journal of Experimental Psychology*, 1971, 87, 369-372.
- Tolman, E. C. *Purposive behavior in animals and men*. New York: Appleton-Century-Crofts, 1932.
- Ursin, H. Effect of amygdaloid lesions on avoidance behavior and visual discrimination. *Experimental Neurology*, 1965, 11, 298-317.
- Ursin, H., & Kaada, B. R. Functional localization within the amygdaloid complex of the cat. *Electroencephalography and Clinical Neurophysiology*, 1960, 12, 1-20.
- Walker, E. L. Reinforcement—"The one ring." In J. T. Tapp (Ed.), *Reinforcement and behavior*. New York: Academic Press, 1969.
- Weiner, B., Heckhausen, H., Meyer, W.-U., & Cook, R. E. Causal ascriptions and achievement behavior: A conceptual analysis of effort and reanalysis of locus of control. *Journal of Personality and Social Psychology*, 1972, 21, 239-248.
- Wessman, A. E., & Ricks, D. F. *Mood and personality*. New York: Holt, Rinehart & Winston, 1966.
- Young, M., Benjamin, B., & Wallis, C. The mortality of widowers. *Lancet*, 1963, 2, 454-456.
- Zeigarnik, B. Ueber das Behalten von erledigten und unerledigten Handlungen. *Psychologische Forschung*, 1927, 9, 1-85.

(Received May 6, 1974)

Toward a Psychological Theory of Alienation

Daniel Stokols
University of California, Irvine

The experience of alienation is conceptualized as a sequential-developmental process which (a) develops in the context of an ongoing relationship between an individual and another person or group of people; (b) involves an unexpected deterioration in the quality of outcomes provided to the individual by the other(s); and (c) persists to the extent that the individual and the other(s) remain spatially or psychologically proximal. On the basis of two variable components, personal-neutral thwarting and salience-nonsalience of alternatives, a typology of estrangement experiences is proposed which includes four behavioral syndromes: isolation, reintegration, subjugation, and rebellion. Though in each of these syndromes the individual's appraisal of the situation leads to at least temporary disillusionment, the induction of persisting cognitive changes as a result of one's exposure to personal thwarting by another is viewed to be the central and unique feature of the alienation experience. Finally, research directions and policy implications of the proposed framework are discussed.

Although the term "alienation" has been associated with numerous connotations and used to describe a variety of phenomena, it nonetheless evokes specific images in relation to contemporary American culture. The ghetto riots and student demonstrations of the sixties, the hippie movement, the drug culture, and the recent rash of political kidnappings—all of these phenomena come to mind as manifestations of alienation; and all seem to involve segments of the population who are, in some sense, "alienated."

This manuscript is based on a doctoral dissertation submitted by the author to the Department of Psychology, University of North Carolina at Chapel Hill. Preparation of the manuscript was partially supported by National Institute of Mental Health Predoctoral Fellowship 1 F01 MH 53585 01 and by research grants from the Carolina Population Center in Chapel Hill and the Program in Social Ecology at the University of California, Irvine. Special thanks are extended to John Thibaut, the author's dissertation advisor, whose insight greatly contributed to the development of the ideas presented here. Critical readings of an earlier version of this manuscript by the other members of the dissertation committee, John Schopler, Robert Mayer, James Wiggins, and Stephen Worchel, are deeply appreciated, as are the helpful comments of Nehemia Friedland, Jeanne Stokols, Gilbert Geis, Arnold Binder, and John Monahan.

Requests for reprints should be sent to Daniel Stokols, Program in Social Ecology, University of California, Irvine, California 92664.

Despite the fact that alienation represents a pervasive theme in American culture and connotes common meanings to many people, our understanding of the determinants and manifestations of alienation remains quite limited. Numerous conceptualizations of alienation have been proposed from historical, philosophical, sociological, and psychological perspectives; yet, very little empirical research concerning the phenomenology of alienation heretofore has been conducted.

The lack of such research seems largely attributable to the nature of previous theoretical perspectives. In general, alienation has been conceptualized as a static phenomenon rather than a dynamic one. Historical and sociological analyses (e.g., Bell, 1960; Blauner, 1964; Durkheim, 1897/1951; Marx, 1844/1964; Merton, 1938; Simmel, 1902/1950; Weber, 1958) typically focus upon social-structural factors which appear to be linked to conditions of widespread alienation within society; for example, automation, anonymity, bureaucratization, and economic inequality. Psychological and philosophical analyses (e.g., Camus, 1956; Keniston, 1965; Kierkegaard, 1959; Sartre, 1953; Seeman, 1959) place a greater emphasis upon the individual's experience and expression of alienation, but these conceptualizations also appear to be static ones;

alienation is viewed either as a personality disposition or an a priori condition of human existence.

The conceptualization proposed here emphasizes the experiential dimensions of alienation but extends earlier psychological approaches in attempting to identify and operationalize the determinants and core dynamic of this experience. The analysis incorporates three fundamental components: (a) a set of antecedent conditions, deriving from one's physical/social environment, which engenders (b) a specific psychological experience, having motivational overtones, and expressed as (c) a set of behavioral manifestations. Alienation is viewed, then, as a sequential-developmental process through which certain social and physical conditions within a particular environment evoke specific feelings and behaviors in its occupants.

The present discussion of alienation is focused primarily at the small-group level. Considering that previous analyses generally have been developed within a broader social context and, thus, have emphasized the relationship between sociocultural conditions and alienation, the small-group setting appears to be an atypical one in which to examine the phenomenon of alienation. The microcosm of the small group, however, offers an advantageous context in which to study alienation for it permits a dynamic, situational analysis and thereby facilitates the development of predictive theory from which experimental hypotheses can be derived and tested. An individual's relationship with society is of an almost abstract nature and certainly of great duration. By contrast, his/her interpersonal relationships which involve specific individuals are often initiated and terminated within a relatively short period of time. Therefore, it becomes more feasible to examine the onset and unfolding of one's alienation from other people rather than from society, since the temporal stages of alienation are more readily detected and delimited at the former level of analysis than at the latter.

An important assumption underlying the proposed theoretical framework is that there

are intrinsic similarities between interpersonal and sociocultural alienation, and that an investigation of the former has relevance for an understanding of the latter. This assumption in no way implies that the experience of alienation as it occurs within small-group settings is fully equivalent with the varieties of sociopolitical alienation discussed by Marx (1844/1964), Durkheim (1897/1951), and Merton (1938). Rather, it suggests that the psychological experience of alienation, in general, involves basic process features which are reflected at both interpersonal and societal levels of analysis.

Specifically, it is assumed that an individual's alienation develops within the context of an ongoing relationship between himself and some other entity—a person, group, society, or culture, for instance. The experience of alienation is brought about through a decline in the quality of one's relationship with a particular context, and this perceived deterioration evokes dissatisfaction with the present situation and a yearning for something better which has been either lost or, as yet, unattained.

The dynamic of disillusionment, suggested above, seems evident in earlier conceptualizations of alienation. In the Marxian analysis, the attainment of class consciousness among proletarian groups reflects a historical process in which objective conditions of exploitation are apprehended by certain segments of society and amended through protest and revolution.¹ Similarly, Merton's

¹ In the writings of Marx, alienation is discussed from both philosophical and ideological perspectives. In a philosophical context, alienation is understood in terms of social and historical forces which eventuate in the objectification of labor (i.e., the spiritual and physical separation of labor from its products; cf. Marx, 1844/1964). From an ideological perspective, conscious recognition of the objectification and exploitation of labor provides the basis on which class consciousness is attained and the forces of revolution are mobilized (cf. Marx & Engels, 1848). It is the ideological dimension of alienation which appears to imply a disillusionment process, though it should be noted that the Marxian analysis pertains to collective consciousness whereas the present one focuses upon the experience of the individual. It seems reasonable to assume that the attainment of class consciousness presupposes the shared estrangement of

portrayal of American society in terms of its "structural inconsistency" depicts a situation in which disadvantaged groups resort to violence as a means of achieving a more equitable distribution of material goods. Moreover, within philosophical and psychological analyses of alienation, a disillusionment process is reflected in the existentialist's feeling of abandonment, discussed by Sartre (1953), and the experiences of "Pyrrhic victory" and "loss of Eden" among alienated college students, as described by Keniston (1965).

The present analysis, then, begins with the assumption that the experience of alienation essentially involves a disillusionment process. From this perspective, Thibaut and Kelley's (1959) comparison level theory is particularly useful as a conceptual base from which to move toward an operational model of alienation. Comparison level denotes a criterion of outcome acceptability with which an individual evaluates the attractiveness of a situation in terms of what he expects or feels he deserves. Assuming that alienation involves a disconfirmation of expectations regarding the quality of a relationship, the comparison level construct is germane to a consideration of the circumstances under which disillusionment occurs and the behavioral concomitants of this experience.

While previous conceptualizations of alienation share certain continuities, their dissimilarities are readily apparent. First, each analysis reflects a characteristic concern with either a societal/cultural or personal/social context of alienation. Historical and sociological perspectives typically focus on the relations among groups within society, whereas psychological and philosophical approaches generally examine the individual's experience in relation to other persons.

A second dissimilarity among earlier analyses concerns the type of disillusionment and behavioral response suggested by each. The proletarian revolutions predicted by

individual workers from the capitalist system of production. Reference to Marx in this discussion, though, is in no way intended to equate sociological and psychological notions of alienation, but merely to suggest a possible parallel between elements of disillusionment in both.

Marx (1844/1964), and the structural inconsistency discussed by Merton (1938), exemplify situations in which segments of society come to recognize their inferior or disadvantaged position vis-à-vis other groups. To the extent that such conditions are prolonged and the sources of inequity are identifiable, the estrangement of the downtrodden tends to intensify and provoke angry protest. In comparison with the analyses of Marx, Durkheim, and Merton, those of Keniston, Sartre, Simmel, and Weber focus upon a qualitatively different set of alienating social conditions. The alienation discussed by the latter group appears to be more passive and less hostile in nature and results from a generalized dissatisfaction with one's milieu. The sources of such dissatisfaction are abstract and impersonal rather than specific and intentional.

To account for the divergent patterns of alienation suggested in previous discussions, the present analysis develops an extension of comparison level theory which incorporates the dimensions of personal-neutral thwarting and salience-nonsalience of alternatives. The interaction of these components suggests a typology of behavioral patterns which seems consistent with earlier analyses.

It should be emphasized that the theoretical perspective developed in the ensuing paragraphs is concerned specifically with interpersonal phenomena. The concepts introduced, by providing a possible link between the psychological experience of individuals and their collective action in response to mutually alienating conditions, may prove to be applicable to an analysis of intergroup relations—the traditional concern of sociological perspectives on alienation. Explicit consideration of the interface between interpersonal and intergroup phenomena, however, is beyond the purview of this discussion.

A THEORETICAL FRAMEWORK FOR THE STUDY OF ALIENATION IN THE SMALL GROUP

The analysis begins at a general level and subsequently turns to a more specific consideration of the small-group situation. The

initial stages of the analysis provide a recapitulation of certain ideas introduced earlier, whereas the latter stages develop those ideas more fully in terms of social psychological theory.

As noted above, the proposed conceptualization of alienation includes three basic components: a constellation of antecedent conditions, a specific psychological experience, and a set of behavioral manifestations. Each component includes a unique set of dimensions. It is useful to focus first upon the psychological experience of alienation since an analysis of this component leads quite naturally to a consideration of its antecedents and consequences.

The experience of alienation. Alienation can be viewed as a form of dissatisfaction resulting from one's perceived association with a negatively valued activity, person, group, or culture. This preliminary definition suggests two important aspects of the alienation experience. First, it develops with reference to some particular context in which a person, P, feels embedded and toward which he feels negative sentiment. The spatial scale of the context may vary along a continuum ranging from the personal, to the social and cultural levels. Second, the experience of alienation persists over time and remains salient to P as long as he feels constrained to the undesirable context. As P becomes aware of his simultaneous proximity to, yet dislike of, a particular contextual referent, his recognition of this discrepancy will lead to frustration. The persistence of frustration, arising from one's constraint to an unsatisfactory context, should result in specific motivational overtones—for example, the desire to dissociate oneself from the context, the inclination to search for alternatives, or tendency to injure some person or object.

It is evident that the above definition does not permit a clear distinction between alienation and other psychological phenomena. The conditions specified would be applicable, for example, to the general experience of dissatisfaction as described by Thibaut and Kelley (1959), psychological reactance as discussed by Brehm (1966), or cognitive imbalance as portrayed by Heider (1958).

In order to distinguish alienation from other theoretical constructs, its spatial and temporal dimensions must be considered more fully.

The antecedents of alienation. Within the personal, social, and cultural contexts of alienation, a common configuration of elements is found: (a) an object, O, with which P feels associated; (b) a type of thwarting, with respect to a given set of outcomes, having a particular source, direction, and either high or low intentionality; the thwarting, whether emanating from O or not, leads to either a temporary or a permanent devaluation of O by P; and (c) a set of forces which constrains P to his association with O; these forces may result directly from O's actions toward P, or through the unavailability of better alternative relationships.

The object of P's alienation, within a personal context, may be represented as an activity in which P is engaged. The worker, for example, becomes alienated from his work as a result of its monotony and impersonality. Within a cultural context, P may perceive himself to be associated with a particular country through formal citizenship yet extremely dissatisfied with certain aspects of citizenship, such as having to serve in the army and participate in an unjust war. Finally, in the context of a small-group situation, P's alienation may derive from an inequitable distribution of opportunities for participation, or certain pre-interaction constraints (e.g., group size) which impair his integration into the decision-making process. Although P is unhappy with his present situation, he is forced, for lack of a better alternative, to remain associated with or proximal to the group.

In terms of Heider's (1946, 1958) balance theory, each of the above examples reflects a similar configuration of circumstances in which P's unit and sentiment relations with a particular object, O, are discrepant. According to Heider, P's relationship with O is balanced to the degree that his perceived linkage with O coincides with favorable feelings toward O (or conversely, P's removal from O is consistent

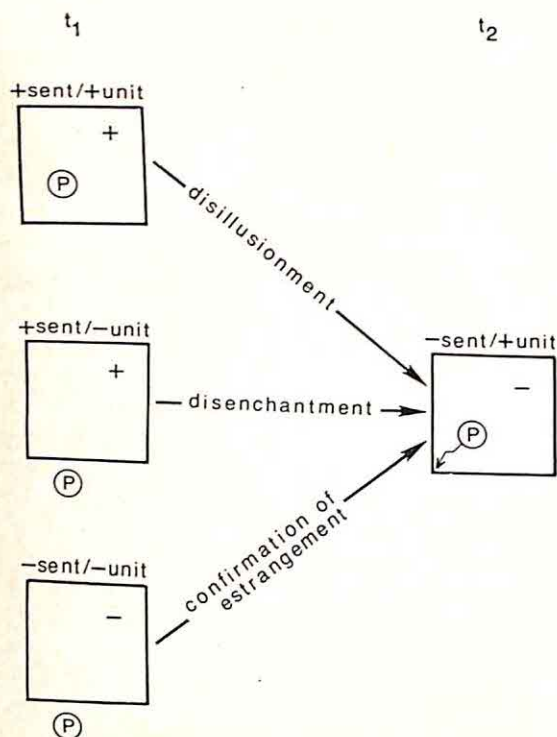


FIGURE 1. Sentiment/unit combinations and the sequential patterns of alienation. (The inclusion of P, an individual, in the square denotes a unit relation, whereas P's presence outside the box designates the absence of one. The plus or minus sign within the square denotes P's sentiment toward the context or object, O. The arrows represent the developmental paths leading to the configuration of alienation.)

with a negative evaluation of O). It follows, then, that in the three examples discussed above, a state of imbalance between P and O exists. Yet, the characterization of alienation as a simple state of imbalance sheds little light upon the specific ways in which the experience develops, the parameters of its intensity, or the form of its manifestation. To gain some understanding of these issues, the temporal, or developmental, nature of alienation must be explored further.

Considering P's relationship to O, four combinations of sentiment/unit relations are possible:

1. positive sentiment/positive unit,
2. negative sentiment/negative unit,
3. positive sentiment/negative unit,
4. negative sentiment/positive unit.

In Figure 1, each of these combinations is represented in diagrammatic fashion. In the present analysis, it is proposed that the fourth possibility, namely, the negative sentiment/positive unit combination, describes the pattern of circumstances reflected in a situation of alienation. Furthermore, each of the other three combinations denotes an initial state of affairs from which the configuration of alienation may be reached. Three developmental patterns or paths, leading to the conditions of alienation, are posited: (a) "disillusionment" or reversal, (b) "disenchantment," and (c) "confirmed estrangement." The path traveled depends upon which of the three initial sets of sentiment/unit relations precedes the configuration of alienation ($-sent/+unit$).

Each of the four possible sentiment/unit combinations can be characterized with respect to two dimensions, namely, the degree of P's *involvement* in his relationship with O, and P's *expectation* of either positive or negative outcomes from his association with O. In the context of this discussion, the dimension of involvement refers specifically to those variables which concern the amount and immediacy of P's experience with O (e.g., the duration of the P-O relationship, or P's spatial proximity to O), rather than P's subjective orientation toward O (e.g., his commitment or enthusiasm toward O). Expectation, on the other hand, relates to the quality and stability of outcomes anticipated by P on the basis of his experience with O. From the above, it follows that close and continuous involvement with O could eventuate in either positive or negative expectations, depending upon the quality of P's cumulative experience with O. Moreover, little or no direct contact with O might set the stage for either positive or negative expectations, depending upon the information P receives about O through others.

In the $+sent/+unit$ situation, P's involvement and expectations are both high. This pattern is reflected, for example, in the case where one has been an active member of a group and has experienced consistently favorable outcomes throughout the term of

his membership. The +sent/-unit situation reflects low involvement but high expectations on the part of P regarding his potential relationship with O. The case in which a freshman attends the rush of a highly desirable fraternity is an example of this pattern. The -sent/-unit situation involves low involvement and low expectations; P is neither related to O at present nor does he wish to become associated with O in the future. Finally, the -sent/+unit pattern of alienation reflects high involvement and low expectancies for favorable outcomes, on the basis of P's present dissatisfaction with O.

It is apparent that the dimensions of involvement and expectation provide a basis for considering the determinants of the intensity of alienation and the form of its manifestation over time. From an extension of Thibaut and Kelley's (1959) comparison level theory, it is expected that P will experience the greatest degree of comparison level discrepancy when his experience of alienation arises through the process of disillusionment or reversal. For, it is along this path that P begins with high involvement and expectations only to arrive at, or be jolted into, an extremely unfavorable situation. To determine why this turn of events sets the stage for the most intense experience of alienation, the relationship between the dimensions of involvement and expectations must be considered more fully.

Any situation in which P has experienced consistently favorable outcomes from O should promote the development of P's commitment to O (cf. Horwitz, 1954; Thibaut, Friedland, & Walker, 1974). That is, P comes to "believe in" O as a reliable source of positive outcomes. Somewhat related to this proposition is Jones and Gerard's (1967) concept of "basic antinomy" which posits that an individual's attributions about a particular entity are most amenable to change during the early stages of exposure to the entity but become progressively more inflexible as the duration of exposure or interaction increases. Thus, as a favorable relationship continues, P should become increasingly committed to O and resistant to

any information which might disconfirm his belief in O. However, once the evidence of O's sudden inadequacy or betrayal exceeds P's "threshold of disconfirmation," the resulting dissolution of P's commitment and expectations should induce a level of disillusionment whose intensity is magnified by the extent of P's prior involvement with O.

In the process of disenchantment, P begins with high expectations but with a minimal amount of previous exposure and commitment to O; and subsequently finds himself in a worse-than-expected relationship. The freshman who has been invited to join the initially desirable fraternity realizes his mistake and entertains the thought of depledging. However, as yet he feels unable to do so because of certain social pressures. This situation provides the basis for a moderate degree of comparison level discrepancy and accompanying alienation, since the reversal of P's expectations does not also involve an obliteration of earlier commitment to O.

The pattern of confirmed estrangement creates the lowest level of comparison level discrepancy for P. He begins with minimal expectations and prior involvement, only to have his predictions confirmed through unavoidable association with O. Consider, for example, the black student living in a predominantly white school district. He is forced by his residence to attend a white school. If he expects, prior to his enrollment, to encounter racism and prejudice at the school and later finds his predictions to be accurate, he will experience alienation. However, it is expected that the intensity of his experience will be less than what it could have been had he entered the school anticipating favorable outcomes.

The involvement and expectation dimensions offer only a partial basis for predicting the intensity of alienation. Other factors may interact with these dimensions to modify their impact on P. Specifically, more information is required concerning the factors which instigate the processes of disillusionment, disenchantment, and confirmed estrangement. A consideration of those thwartings which lead to a deterioration in

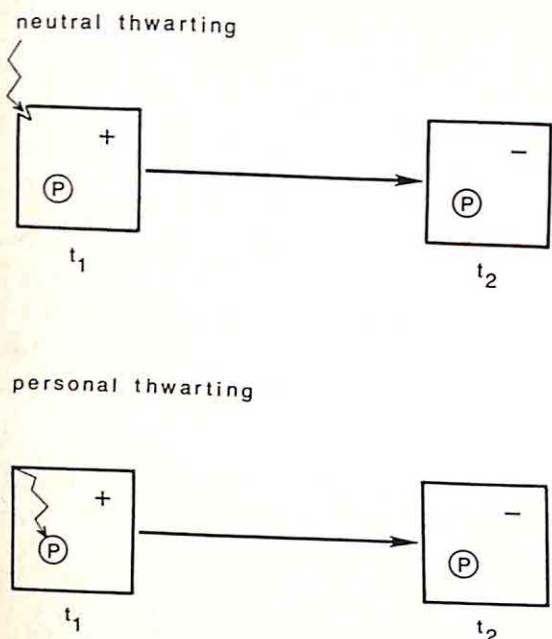


FIGURE 2. Neutral and personal processes of disillusionment. (The inclusion of P, an individual, in the square denotes a unit relation whereas the presence of P outside the box designates the absence of one. The plus or minus sign within the square denotes P's sentiment toward the object or context, O. The arrows represent the developmental paths leading to the configuration of alienation.)

P's relations with O should provide the needed information.

The remaining discussion will focus upon the process of disillusionment for two reasons. First, it is necessary to assume that a relationship exists between P and O in order to examine the impact of various types of thwarting on that relationship. This assumption is possible only with respect to the disillusionment pattern of alienation, in which a formerly satisfying relationship with O (even one which only recently has been established) turns sour. Second, the reversal of P's feelings toward O from commitment to disillusionment represents the most typical characterization of alienation, as described in the historical, sociological, philosophical, and psychological literature.

As stated earlier, P's experience of alienation arises from a specific thwarting which strains his relationship with O. A thwarting may be characterized with respect to three basic dimensions: (a) its *source*, (b)

direction, and (c) *intentionality*. The first refers to the origin of the force; that is, whether it emanates from O or from the extraneous environment. The second relates to the target of the force; for example, whether or not it is directed specifically at P. And the third concerns P's attribution of intentionality to the source of the force. These dimensions may be combined in a variety of ways to yield different types of thwartings.

The psychological impact of a particular thwarting will be mediated by many contextual variables. A highly directed and intentional thwarting, for instance, could be interpreted differently by P depending upon the information he has concerning the norms or situational pressures which explain O's actions. When O's actions are interpreted as reasonable in view of the particular situation (e.g., when it becomes necessary for P to accept an undesirable role for the good of the group), less dissatisfaction and alienation should be aroused in P than when O's actions appear to be arbitrary and unjustified (e.g., P is unfairly chastised or insulted by O). In a similar vein, an undirected and unintentional thwarting might arouse divergent emotions in P depending upon his prior experience with O. To the extent that the relationship has been unfavorable in the past, P may assimilate even an unintentional thwarting to the characteristic deficiencies he ascribes to O.

Two types of disillusionment processes are represented in Figure 2. Each results from a different kind of thwarting. The first case involves a *neutral thwarting*—that is, one which does not emanate directly from O, is not specifically directed at P, and is perceived by P as being unintentional. Spatial constraints, information overload, and large group size represent neutral thwartings at the small-group level. Each constraint hinders P's relationship with O in some sense. Assuming, for example, that P desires to participate in group discussion and decision making, this need may be thwarted by environmental circumstances which make it difficult for him to interact meaningfully with the group. Lack of avail-

able space impairs the coordination of P's activity with that of other group members; and the presence of too many members reduces P's opportunity to participate in group decision making.

The second type of disillusionment process involves a *personal thwarting*—that is, one which stems directly from O, is specifically directed at P, and is perceived by P as reflecting O's intentions. A hostile or inequitable action represents an example of a personal thwarting. The intentional relegation of P to tedious and low-status roles is likely to be perceived by him as unfair discrimination, exclusion, and rejection by the group.

In view of the differences between the neutral and personal patterns, it is expected that P's disillusionment will be more bitter and irreversible in the latter case than in the former one. In both processes, P experiences a reversal in his feelings from high commitment and expectations to unanticipated frustration. As suggested earlier, such reversal results in comparison level discrepancy and puts a strain on P's commitment to O. In the personal pattern, though, the source of P's frustration is more specific, and its intentions more visible and malevolent. The sudden communication of unnecessary insult or injury by O introduces the element of "rejection" into P's experience of disillusionment. Thus, the "gain-loss" element (Aronson & Linder, 1965) arising from P's rejection by O, when coupled with the frustration of P's expectancies regarding favorable outcomes, should promote a more intense and irreversible experience of alienation than when comparison level discrepancy operates alone. Moreover, rejection, by definition, implies a low correspondence between the outcomes of P and O (Thibaut & Kelley, 1959); as well, there is a tendency on the part of P to reciprocate O's rejection (Berscheid & Walster, 1969). Hence, the possibilities for salvaging a satisfactory relationship with O are minimal, and the probability that P will dissolve his commitment to O is high.

In the neutral process of disillusionment, however, the source of P's frustration is

vague and less personal in nature. Because there is no ready source to which P can impute foul play, the commonality of interest between P and O is greater than in the personal case. And because there is more basis for cooperation in the neutral case than in the personal one, P will be more amenable to an eventual reconciliation with O in the former instance. Thus, although a neutral thwarting may at first seem unjustifiable to P in view of O's previous reliability as a source of positive outcomes, the strain in P's commitment to O should soon dissipate as long as evidence of O's malevolent intent continues to be lacking.

The term, "thwarting," has been interpreted in the present analysis as a force which provokes a reversal in the quality of P's relationship with some object, O. The reversal has been characterized as a progression from a satisfactory state of affairs, endorsed by P, to an unsatisfactory situation involving his disillusionment. Hence, the term thwarting, as used above, refers to a type of force which provokes P's experience of alienation and partially determines the intensity of that experience.

In order to arrive at a more adequate representation of the alienation experience, however, we must consider the impact of a particular thwarting in the context of those "restraining forces" which enforce P's continued association with O once the reversal process has occurred. For, without P's prolonged proximity to or association with the object of his disillusionment, there would be no basis for his experience of alienation. He would merely dissociate himself from the alienating context, and its aversive salience would thereby decrease.

The $-sent/+unit$ configuration of alienation, thus, reflects a "nonvoluntary" relationship (Thibaut & Kelley, 1959). In such situations, P is forced, over time, to remain at an infra-comparison level state of outcome quality. A variety of restraining forces may operate to enforce P's association with the object of his alienation; physical pressures emanating directly from O, or P's lack of more desirable alternative relationships, for example. In the small-group

situation, P may be forced by social pressures against "copping out," to continue his membership with some group; or he may decide to continue his association with the group for want of a better alternative. In either case, P perceives that desirable alternative relationships are unattainable and consequently remains in an unsatisfactory situation.

The behavioral manifestations of alienation. The behavioral component of alienation can be considered in terms of its *form* and its *effects*. The form of response to felt alienation can be characterized as aggressive or nonaggressive, active or passive, and directed or nondirected. The hostility, intensity, and directionality of a person's behavior would convey the quality and strength of his feelings. Regarding its effects, behavior can be viewed as either adaptive or maladaptive; that is, as an action which alleviates or intensifies P's experience of alienation.

It is proposed that P's behavior will be most active, directed, and aggressive in response to a personal thwarting, where the source of one's frustration is specific and intentional. Furthermore, the gain-loss element provokes a clear conflict of interest between P and O; hence, minimal opportunity exists for salvaging the damaged relationship. This situation, then, which elicits frustration and provides cues for hostility (manifested in O's rejection of P), as well as a specific target for P's counteraction, is highly conducive to aggressive behavior (cf. Berkowitz, 1965; Buss, 1961).

The neutral pattern of reversal, by contrast, provides a greater range of behavioral options for P. Since the event leading to P's frustration is unintentional and its source vague, a basis for cooperation between P and O exists. Thus, P's response to a neutral thwarting will be of a problem-solving nature. His behavior may assume a variety of forms over time. In response to spatial constraints, for example, P may initially attempt to improve the coordination of his activity with that of the group. If his efforts are ineffective, he may withdraw temporarily into passive isolation. Pro-

longed frustration, however, may eventually lead to some form of hostility or permanent isolation.

The form of P's response to a particular thwarting will also be influenced by the salience of desirable alternative relationships. The present conceptualization presupposes the existence of a positive unit relation between P and O. That is, P perceives himself to be linked to O through spatial proximity or the expectation of future proximity. Moreover, this perceived linkage is unpleasant and nonvoluntary to the degree that P finds O unattractive and presently feels unable to improve their relationship or establish an alternative one. Although for the time being he is temporarily and unwillingly constrained to O, it is quite possible that he may expect to achieve a more satisfactory relationship in the near future. Regardless of whether viable alternatives are actually available to P, the anticipation of an improved relationship with O or an available substitute should markedly affect the intensity and overt aggressiveness of P's response to either a personal or neutral thwarting.

P's behavior may be viewed in the context of four situational configurations:

1. neutral thwarting/perceived lack of alternatives,
2. neutral thwarting/salient alternatives,
3. personal thwarting/perceived lack of alternatives,
4. personal thwarting/salient alternatives.

The combinations are ordered with respect to the degree of frustration each evokes in P. The perception of available alternatives in the context of a neutral thwarting should engender the least amount of frustration, whereas a personal thwarting, in the face of salient alternatives, would evoke the most. Hence, the instigation to aggression should be strongest in the latter situation and weakest in the former.

Four ideal types of response, to the various combinations of restraints, are suggested:

1. neutral thwarting/perceived lack of alternatives → isolation
2. neutral thwarting/salient alternatives → reintegration
3. personal thwarting/perceived lack of alternatives → subjugation
4. personal thwarting/salient alternatives → rebellion

Assuming that P is, for the time being, constrained to an unsatisfactory relationship with O, either by force or current lack of alternative situations, what are the behavioral options open to P in each of the above configurations? Thibaut and Kelley (1959) propose two means by which an individual can reduce his frustration in a nonvoluntary relationship: (a) lower the comparison level, or (b) find some way to achieve the presently unattained level of outcome quality. Each behavior represents an adaptive response to felt alienation.

The lowering of one's standards regarding the quality of outcomes he deserves would be most likely to occur in situations where desirable alternative relationships were perceived to be unattainable. In contrast, continued search for a way of meeting current standards of outcome quality would be more characteristic of situations in which favorable alternatives were viewed as potentially available.

To predict the degree to which P's behavior will be adaptive in the context of either salience or nonsalience of alternatives, it is necessary to consider further the behavioral implications of personal and neutral thwartings. Where alternatives are perceived to be unavailable, a personal thwarting would make it more difficult for P to accept presently unsatisfactory outcomes than would a neutral thwarting, since the rejection and arbitrary insult conveyed by the former would frustrate P's hope for a future improvement in his relations with O. Thus, the frustration resulting from subjugation to an unreformable situation would be greater than that stemming from temporary isolation in a situation which may subsequently improve.

Similarly, in situations where more desirable alternatives are salient, P's potential

for effective adaptation would be greater in the context of a neutral thwarting, as compared to a personal one. Given that P is currently constrained to O, his range of options in dealing with an unintentional thwarting by O would be greater than those associated with a personal rejection by O. The first situation offers P the possibility of deferring judgment on O and actively attempting to improve the present relationship, before endeavoring to reach external alternatives. A personal thwarting by O, however, reduces P's opportunities for reintegration and increases his potential for frustration resulting from an unrelievable discrepancy between his negative feelings toward, yet proximity to, O. Hence, until imposed proximity with O was somehow reduced, the likelihood of P's rebellion against O in the form of overt protest and aggression would be great.

Recent research literature on social psychological stress (cf. Lazarus, 1966; Levine & Scotch, 1970; McGrath, 1970) would suggest that when P is unable to alleviate his experience of alienation, and this experience extends over a prolonged period of time, two general syndromes of stress may ensue: psychophysiological stress as reflected in certain physical disorders (e.g., Wirth, 1938), and self-destructive or anti-social behavior as manifested in P's self-disparagement or aggression toward others (e.g., Durkheim, 1897/1951; Merton, 1938).

Differentiation of Alienation from Related Psychological Constructs

Having discussed the situational, experiential, and behavioral components of alienation in greater detail, it is now essential to specify the dissimilarities between alienation and other psychological phenomena. An important question in this context is whether or not the proposed conceptualization of alienation is so broad as to embrace any form of disaffection. An excessively general framework would shed little light on the distinctive aspects of alienation.

The phenomenology of alienation, as described earlier, does include elements of dislike, dissatisfaction, imbalance, frustration,

and reactance. As P becomes disillusioned with O, he develops a dislike of O and feels dissatisfied with the inferior (infra-comparison level) quality of his present outcomes. Moreover, P's embeddedness in a negatively valued context should promote cognitive imbalance and frustration, as well as feelings of preemption and infringement which might be characterized as reactance.

Although the experience of alienation subsumes the above-mentioned elements, it can be distinguished from them on the basis of its etiology, psychological impact, and behavioral consequences. According to the present analysis, P's alienation results from a deterioration in his formerly satisfactory relationship with O. Moreover, P expects to remain proximal to O even after the reversal has occurred. This developmental pattern does not apply necessarily to feelings of dislike, dissatisfaction, imbalance, or reactance. P's dislike of or dissatisfaction with O does not presuppose an originally pleasant relationship between himself and O. In the case where P is blindly prejudiced or has extremely high expectations, for example, he may adopt a negative or skeptical orientation toward O, even before meeting him. And the constraint to remain in a negatively valued situation is not a necessary condition for the arousal of reactance. Reactance can occur even when P is forced to do what he wants or say what he wants to say (Worchel & Brehm, 1970).

The present analysis of alienation relies heavily on the distinction between personal and neutral thwartings. While differences between these constructs were discussed earlier, it is important here to emphasize their similarities and uniqueness in relation to other types of thwartings for purposes of conceptual clarification. The notion of thwarting implies the arousal of frustration in P resulting from an interference in the gratification of a previously aroused motive (Cofer & Appley, 1964; Dollard, Doob, Miller, Mowrer, & Sears, 1939). The context in which frustration occurs can vary from transitory encounters to ongoing relationships. Both personal and neutral thwartings, as defined here, however, occur

within a previously pleasant relationship and, as such, both promote some degree of disillusionment in P with regard to O. P's disillusionment derives from his prior commitment to O and the unexpected but suddenly apparent possibility that his commitment may no longer be tenable.

The unexpectedness of personal and neutral thwartings, in view of O's previous reliability as a source of positive outcomes, should lead P to perceive either type of reversal as unjustifiable. The unjustifiable aspect of these thwartings renders them conceptually similar to Pastore's (1952) notion of arbitrary frustration, though the latter construct has typically been confined to personally, rather than impersonally, delivered thwartings (cf. Burnstein & Worchel, 1962; Cohen, 1955; Rothaus & Worchel, 1960). While both types of thwartings should arouse frustration, the personal process should initiate more hostility due to the evidence of O's malevolence and the insult this conveys to P (cf. Buss, 1961). Thus, personal and neutral thwartings can be identified, among other types of thwartings, as involving an element of perceived unjustifiability, and in the personal case, as implying O's insult of P.

The present emphasis on personal and neutral thwartings is based upon widespread reference to the element of disillusionment, above and beyond frustration and hostility, found in most discussions of alienation. As noted earlier, disillusionment involves a strain in P's commitment to O. Whether or not this strain leads to a dissolution of commitment depends upon P's assessment of the future viability of his relationship with O.

The outcome of P's appraisal will depend largely on the presence or absence of cues which signal the future unreliability of O and an unresolvable "lack of fit" between the attributes, interests, and goals of O and P. In this regard, Etzioni (1968) characterized alienation as a social situation which is "beyond the control of the actor and unresponsive to his basic needs" (p. 879). It seems reasonable to assume that the clearest forewarning of O's continuing unresponsive-

ness or malevolence would be communicated by personal thwartings which pose a threat to P's physical safety, ideology, or self-concept, and imply an irreducible conflict of interest between him and O. On the other hand, the least amount of information suggesting O's future unreliability would be provided by neutral thwartings which are basically unintentional, nondirected, and minimally threatening to P's security.

In situations of neutral thwarting, the ambiguity of O's future orientation would necessitate the adoption of an exploratory set on P's part. That is, he would tend to withhold judgment regarding the viability of his relationship with O until additional, relevant information was obtained. P's search for more information might take the form of active protest, quiet observation, or a pattern involving both elements of protest and contemplation. At any rate, as long as P remained in an exploratory set, his disillusionment with O would not eventuate in a final dissolution of commitment to O.

Under conditions of personal thwarting, however, the jeopardization of P's safety or self-concept would provide clearcut evidence of O's potential unreliability. Such evidence should ultimately provoke P's decision that the psychological distance between himself and O is insurmountable. Once this decision has been reached, P's commitment to O should be withdrawn, his perceptions of O altered, and in some cases, his self-concept modified (as reflected, for example, in P's generalized reluctance to approach others similar to O). These cognitive changes in P are likely to be accompanied by some form of angry retaliation against O for his original insult of P. Because of the postthwarting undesirability of any material or benefits previously provided by O, P's aggression would remain noninstrumental (Buss, 1963) or expressive in nature, and as such (because of its minimal reinforcement value) should dissipate subsequent to an initial display of hostility.

The foregoing discussion provides not only a delineation of personal and neutral thwartings in the context of frustration-aggression theory but also a sketch of the

distinctive experiential features of the alienation syndrome. Three basic stages of the alienation experience can now be summarized: (a) the *sense of disillusionment* resulting from an unexpected thwarting which strains P's commitment to O; (b) a post-thwarting *process of appraisal* by which P evaluates the future viability of his relationship with O on the basis of circumstantial information—this stage may involve instrumental aggression by P aimed at clarifying or rectifying the unsatisfactory situation; and (c) P's *decision that his relationship with O is no longer tenable*—this decision will result in *irreversible cognitive changes* in P which may be manifested through sporadic bursts of aggression.

Whereas stages of disillusionment and appraisal are associated with neutral as well as personal thwartings, the induction of irreversible cognitive changes in P is viewed in this analysis as the unique and central feature of the alienation experience. All instances of alienation would seem to involve the general perception that one's self-image or identity is highly discrepant with others' perception of him. This realization could provide the foundation for any number of attitudinal changes regarding others (i.e., "They are unfair" or "irrelevant") as well as oneself ("I don't need them," "I don't enjoy the activities they do," "My values are basically different from theirs," etc.). To the extent that P's disillusionment with O effects attitudinal changes along several interrelated dimensions, P's alienation from O should become increasingly irreversible over time.

Referring to alienation as an irreversible syndrome is not to imply that cognitive changes associated with disillusionment are totally resistant to modification. Indeed, reliable evidence of O's postthwarting reforms may well promote a rapprochement between P and O at some time in the future. Rather, the description of alienation, as irreversible is meant to distinguish its phenomenological features from those of other psychological phenomena which are relatively more transitory in nature (e.g., momentary frustration or dissatisfaction).

The main contention here, then, is that the cognitive changes associated with P's estrangement from O extend (both in a qualitative and temporal sense) beyond feelings of frustration, imbalance, dissatisfaction, and dislike, and are not predicted on the basis of theoretical constructs pertaining to these feelings.

The relative irreversibility of the alienation experience is perhaps its most distinctive quality vis-à-vis similar psychological phenomena. Many theories describe adaptive mechanisms which enable P to reconcile his relationship with an initially unfavorable entity. From the perspective of balance theory (Heider, 1958), for instance, P's perception of a unit relation between himself and O should promote positive sentiment toward O over time. Similarly, cognitive consistency theory posits processes of denial, differentiation, and transcendence (Abelson, 1959) which operate to resolve the discrepancies in P's sentiment and unit relations with O. Thibaut and Kelley (1959), in their discussion of nonvoluntary relationships, suggest that P may reduce his dissatisfaction with O by lowering the comparison level. And Dollard et al. (1939) propose that frustration can be reduced through cathartic aggression. The present analysis, however, contends that P's alienation, when brought about through a personal thwarting by O, will tend to remain unresolved as long as he remains proximal to O, even after the objective frustrations (e.g., loss of status, salary) associated with the original thwarting have been removed. Only by withdrawing from or obliterating his relationship with O can he reduce the weight and salience of his acknowledged separateness, unengagement, and estrangement.

As indicated earlier, the potential for P's reconciliation with O will be considerably greater when his alienation results from a neutral thwarting. In view of the temporary and reversible nature of "neutral" alienation, it is perhaps more fruitful to characterize this experience as a transitory state of disaffection rather than a form of alienation. In certain respects, such disaffection is qualitatively closer to the experi-

ence of psychological reactance (Brehm, 1966) than to that of alienation. According to Brehm, reactance is evoked in P when his freedom is directly or implicitly eliminated and prompts efforts on his part to reestablish the lost or threatened freedom. In the present context, P's experience of a neutral thwarting provokes dissatisfaction with O, largely because he is no longer free to enjoy the rewarding outcomes which previously emanated from O. He is motivated, though, to explore possibilities for improving his relationship with O as a means of regaining his freedom to enjoy satisfactory outcomes.

When P's alienation results from a personal thwarting, however, the importance of receiving positive reinforcement from O is diminished. Despite the absence of favorable alternative relationships, P is presumably more willing to surrender his freedom to enjoy rewarding outcomes than to attempt a satisfactory reconciliation with O. Thus, although the experience of "personal" alienation involves feelings of reactance and the motivation to recover the quality of one's outcomes, the specific freedom preempted by O may be devalued and eventually replaced by the "freedom to be one's own man." This substitution of freedoms may account for the behavior of alienated individuals who withdraw from or rebel against O, rather than attempting to salvage the relationship.

RESEARCH DIRECTIONS AND POLICY IMPLICATIONS OF THE PROPOSED FRAMEWORK

The conceptualization of alienation presented above provides the basis for developing a predictive theory which can be empirically examined in the context of intra-group conflict situations. An empirically validated theory would have important implications for the formulation of intervention strategies designed to reduce conditions of alienation within the community.

Toward Empirical Assessment of the Proposed Framework

From certain assumptions concerning the situational antecedents, psychological ex-

perience, and behavioral expression of disillusionment, two dimensions were derived which permit a delineation of the spatial and temporal features of the alienation experience: personal-neutral thwarting and salience-nonsalience of alternatives. On the basis of these variables, alienation was conceptualized not only as a particular psychological state or configuration of relations between O and P but also as a chronological process involving a pattern of emotional reactions and behaviors.

A typology of behavioral patterns based upon the thwarting and alternative dimensions was proposed which included four ideal types of response: isolation, reintegration, subjugation, and rebellion. Each pattern was examined in terms of three experiential stages: a postthwarting period of disillusionment, a process of appraisal, and eventual reaffirmation or dissolution of commitment. The levels of estrangement and aggression associated with these stages provided a basis for distinguishing among qualitatively different syndromes of response to alienating conditions.

Specific predictions pertaining to the behavioral impact of the thwarting and alternatives dimensions can now be advanced:

1. The intensity and duration of an individual's estrangement will be greater under conditions of personal thwarting, than under those of neutral thwarting.

2. A personally thwarted individual will experience more intense and prolonged hostility than one who has been neutrally thwarted.

3. (a) During the early stages of disillusionment, individuals for whom alternative relationships are salient should exhibit more aggression than those who perceive alternatives to be unavailable. (b) Subsequent appraisal of a thwarting as "neutral" should eventuate in reduced aggression and a higher rate of reintegrative behavior, especially among persons for whom alternatives are salient. (c) Appraisal of a thwarting as "personal," however, should result in a sustained level of hostility and a continuance of aggression, especially among individuals for whom alternatives are salient.

- (d) Following a decision to dissolve one's commitment to O, overt aggression should decline to the extent that prior aggressive actions have been ineffective in reducing proximity with O.

These predictions, which assume P's perceived or actual proximity with O, are represented diagrammatically in Figure 3. As a function of the thwarting and alternatives components, the response variables of estrangement, hostility, and overt aggression have been plotted over time to depict the "behavioral profiles" of isolation, reintegration, subjugation, and rebellion.

The graphic portrayal of each profile suggests that any attempt to assess the validity of the proposed typology of behavioral syndromes would require a series of repeated measurements taken simultaneously along three different dimensions of response. Only a time-series approach would permit a qualitative distinction between temporary states of estrangement, as reflected in neutral thwarting conditions, and the process whereby personally thwarted individuals come to feel irrevocably alienated.

Whereas the behavior associated with neutral forms of estrangement would be more reflective of transitory motivational states (e.g., frustration, reactance), the sustained hostility and estrangement accompanying personal alienation would reflect stable dispositional and cognitive changes in P. The distinction between "motivational" (neutral) and "attitudinal" (personal) forms of estrangement implies that the former would subside following the removal of certain objective thwartings (e.g., low salary, status) while the latter would persist despite subsequent overtures by O to restore P to his original position.

Support for the previously stated predictions comes from a recent experiment (Stokols, 1973). In the context of a "Group Productivity Study," each subject worked with three other individuals (actually experimental accomplices) on a series of tasks. The 2-hour experimental sessions incorporated three basic phases which together were designed to create a reversal in the subject's outcomes.

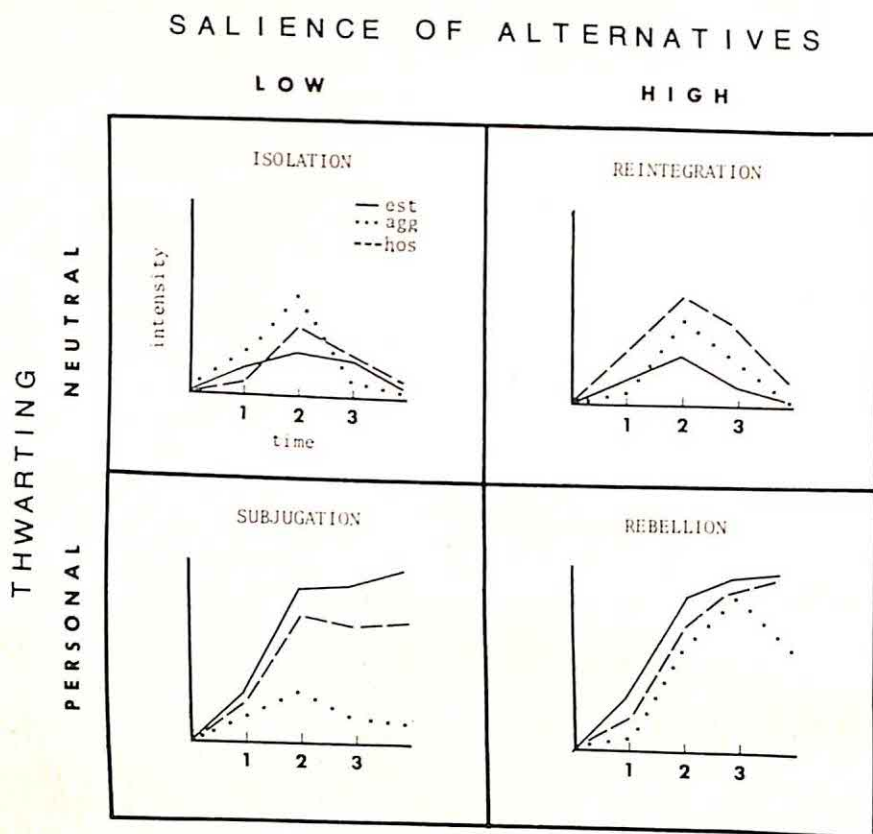


FIGURE 3. A typology of behavioral patterns based upon the dimensions of personal-neutral thwarting, salience-nonsalience of alternatives, and duration of exposure to situation (with perceived proximity to O, an object, held constant). (Times 1, 2, and 3 correspond to the stages of disillusionment, appraisal, and decision. The lines shown in each graph depict the predicted responses of estrangement, aggression, and hostility over time.)

During the first phase, the subject's interaction with the group was rewarding and pleasant; he received positive feedback from the others and, in general, was led to believe that his performance had contributed to the success of the group. In the second phase, though, he experienced a reversal in the quality of his relationship with the group and a reduction in his earnings and status relative to those of the other group members. He was relegated to a mundane and inferior role while the others had access to interesting and enjoyable tasks. During the final phase, the subject found himself in a role which permitted him to exercise control over the other group members; he could choose to ignore, punish, or reward them.

The personal-neutral thwarting dimension was manipulated through the assignment of

the subject to an inferior role during Phase 2 either by a group vote (personal) or random (neutral) process. And the salience or non-salience of alternatives was varied through the assignment of subjects on the basis of self-esteem prescores (obtained prior to the experimental session) to either high-esteem or low-esteem experimental conditions.

The main dependent measures included repeated assessments of hostility, overt aggression, and reported estrangement on a number of Keniston's (Note 1) Short Alienation Scales. Preexperimental measures of chronic alienation and esteem were utilized as covariates in the statistical analyses. Results indicated significantly higher levels of reported estrangement and hostility during Phase 3 among personally versus neutrally thwarted subjects yet the absence of between-groups differences in overt aggres-

sion during the final phase of the experiment. Perhaps most interesting were the changes in subjects' attitudes not only toward other group members but also about themselves—for example, their position on certain ideological issues—as a result of having experienced a personal thwarting.

*Development of Conceptual Linkages
Between Alienation and Other Social
Psychological Phenomena*

The present analysis of alienation is relevant to a consideration of those conditions under which persons are willing to repudiate their prior commitments to individuals or social systems. Most theories of attitude change, especially the consistency theories of Heider (1958), Festinger (1957), and Bem (1967), emphasize the forces which prompt people to remain attitudinally consistent with their previous behavior or public commitments. These formulations, however, generally have not addressed those situations in which individuals refuse to ignore or justify the deterioration of a relationship to which they were previously committed.

Aronson's (1968, 1969, 1972) restatement of dissonance theory suggests that high-esteem persons will feel the most need to justify their behavior when its consequences threaten the integrity of their self-concept or esteem. The present discussion, though, predicts that high- versus low-esteem people, for whom viable alternative relationships are characteristically salient, will be most likely to renounce their association with O, especially when O's actions threaten and insult their self-concept.

An interesting direction for future research would be to explore systematically the conditions under which people will defend or dissolve their earlier commitments in the face of disconfirming evidence regarding the appropriateness, viability, or value of those commitments. The willingness of P to break off with O may well stem from the apparent futility of the situation, as conveyed by the arbitrariness of O's behavior, and the passive (nonvolitional), rather than active (voluntary), exposure of P to O's

insult. In any event, the critical variables in this context can only be ascertained more fully through subsequent research.

Another important direction for future research pertains to the manner in which aspects of the physical and social environment combine to induce stress in the occupants of a particular behavior setting. In the foregoing discussion, various syndromes of extrangement have been viewed as the result of *either* personal *or* neutral thwartings. Interesting extensions of this discussion relate to (a) the circumstances under which impersonal thwartings will be misattributed to personal sources and intentions, and (b) the processes by which personal thwartings become associated with certain aspects of the physical environment, so as to mediate the behavioral and psychological impact of physical stressor variables (e.g., noise, density, pollution).

Assuming that ecological pressures can place strains on the relations among members of a social system, it becomes important to specify the conditions under which neutral thwartings are perceived as personal ones, since such misattribution is likely to result in a more intense, irreversible syndrome of stress than that which would arise from a seemingly unintentional thwarting. Specification of these conditions may prove to be quite valuable in formulating social planning policies amenable to a variety of environmental settings, for example, oversized classrooms and high density neighborhoods.

Moreover, an elucidation of the processes by which physical features of the environment acquire cue value in the context of social thwarting situations should enhance our understanding of human response to environmental stress. For instance, crowding has been characterized as a psychological experience in which an individual's demand for space exceeds the available supply (Stokols, 1972). An interesting parallel can be drawn between this characterization of crowding and the proposed conceptualization of personal alienation: The experience of a personal thwarting should sensitize P to the need for more space as a means of

reducing his proximity to O. Thus, continued exposure to personal thwartings, especially in the context of a high density situation (e.g., a crowded home, classroom, or work environment) would heighten P's sensitivity to density or proximity constraints and lower his resistance against crowding stress; the negative cue value of density might even extend to situations in which O is not present.

Most recent analyses of human crowding do not suggest the circumstances under which specific density constraints (e.g., restriction of movement, stimulus overload, infringement on privacy, or proximity with threatening others) will be more salient than others or the particular psychological and behavioral concomitants of each. An advantage of the personal-neutral thwarting dimension, though, is that it suggests a criterion for determining the intensity, duration, and potential reducibility of various crowding experiences.

On the basis of the proposed thwarting dimension, a distinction can be drawn between "neutral" and "personal" experiences of crowding (Stokols, Note 2). The former variety of crowding would involve increased sensitivity to space as a result of the frustrations associated with neutral constraints (behavioral restriction, stimulus overload). The latter variety of crowding would stem from the security concerns made salient by proximity with hostile or unpredictable persons. Since feelings of crowding in the latter case would relate to the individual's beliefs about the potential dangerousness of others, this type of experience would tend to be more intense, of greater duration, and more generalizable across various situations than would transitory experiences of impersonal density constraints.

Policy Implications

Previous studies have noted the relationship between alienation and a variety of social problems. Forward and Williams (1970), Gurin and Gurin (1970), and Ransford (1968) have discussed the association between feelings of powerlessness and racial violence. Also, in the context of

contemporary concern over the population crisis, Groat and Neal (1967) have examined the relation between social isolation (Seeman, 1959) and the motivation to bear children as "compensation" for feeling isolated.

Despite the recognized relationship between alienation and social problems, little progress has been made toward the development of ameliorative strategies. The lack of such progress may be attributable to the static nature of previous analyses which place a greater emphasis on the measurement of alienation as a personality disposition than upon an analysis of its environmental origins and process features.

Some writers have utilized the concepts of powerlessness and isolation as a guide for developing policies aimed at alleviating alienation among the poor. Gurin and Gurin (1970), for example, suggest that structural changes in the social system, leading to increased opportunities for experiencing success, provide a means of raising the expectancies of low-expectancy people. Similarly, Sarbin (1970), analyzing the "culture of poverty" (Lewis, 1961) in terms of role theory, concludes that society must create opportunities for the poor to realize achieved statuses which permit the exercise of choice, the experience of social praise, and the development of increased self-esteem.

The above recommendations, however, appear to be too general, undifferentiated, and, therefore, of limited value in approaching social problems related to alienation. The increased provision of roles associated with achieved statuses within the system will be ineffective as a means of reducing alienation if the disaffected individuals see themselves as highly competent and powerful in the first place but ideologically unable to participate in any reciprocal relationship with O. Under these circumstances, it may be more fruitful for the system to eliminate cues which signal an enforced relationship between itself and P and to facilitate P's actual or symbolic withdrawal from the relationship as well as his freedom to "do his own thing" outside of the system.

The present analysis suggests that a di-

verse set of interventions may be necessary in dealing with social problems related to alienation. Whether an individual attributes his dissatisfaction with society to personal factors (such as racial discrimination) or to neutral circumstances (e.g., congestion, information overload, and anonymity arising from high densities) will affect the intensity of his alienation, the manner in which it is expressed, and its amenability to ameliorative strategies. Similarly, the salience of desirable alternatives will probably influence his decision either to seek a reconciliation with the system or protest actively and perhaps violently against unsatisfactory social conditions.

Where P has been personally thwarted, the most effective intervention policy from O's point of view would be one which provides evidence contrary to P's attributions. Rather than trying to impose a forced relationship on P, O would attempt to eliminate P's justification for attributing malevolent intent to O. The willingness of City Hall to facilitate community organization and self-government within the ghetto, for instance, would probably serve to reduce the estrangement of its residents.

In situations where P is unable to attribute a deterioration in his outcomes to malevolent forces, he will be more amenable to cooperating with O as a means of achieving a mutually satisfactory resolution of their joint problems. Under these conditions, P will be more receptive to compromise solutions initiated by O. Increased opportunities for participation in the system, then, would be particularly appealing to low-expectancy individuals who previously have been isolated from achieved statuses.

The above speculations are derived from basic assumptions concerning the phenomenology of alienation. These assumptions have been stated throughout this discussion and remain to be validated more fully through future research at both the small-group and sociocultural levels.

REFERENCE NOTES

1. Keniston, K. Alienation scales. Unpublished manuscript, 1965. (Available from the Depart-

ment of Psychiatry, Yale University Medical School, New Haven, Connecticut 06510.)

2. Stokols, D. The experience of crowding in primary and secondary environments. Paper presented at the Annual Meeting of the American Psychological Association, New Orleans, August 1974.

REFERENCES

- Abelson, R. Modes of resolution of belief dilemmas. *Journal of Conflict Resolution*, 1959, 3, 343-352.
- Aronson, E., & Linder, D. Gain and loss of esteem as determinants of interpersonal attractiveness. *Journal of Experimental Social Psychology*, 1965, 1, 156-171.
- Aronson, E. Dissonance theory: Progress and problems. In R. Abelson, E. Aronson, W. McGuire, T. Newcomb, M. Rosenberg, & P. Tannenbaum (Eds.), *Theories of cognitive consistency: A sourcebook*. Chicago: Rand McNally, 1968.
- Aronson, E. Some antecedents of interpersonal attraction. In W. Arnold & D. Levine (Eds.), *Nebraska Symposium on Motivation*, 1969. Lincoln: University of Nebraska Press, 1969.
- Aronson, E. *The social animal*. San Francisco: W. H. Freeman, 1972.
- Bell, D. *The end of ideology*. Glencoe, Ill.: Free Press, 1960.
- Bem, D. Self-perception: An alternative interpretation of cognitive dissonance phenomena. *Psychological Review*, 1967, 74, 183-200.
- Berkowitz, L. The concept of aggressive drive: Some additional considerations. In L. Berkowitz (Ed.), *Advances in experimental social psychology* (Vol. 2). New York: Academic Press, 1965.
- Berscheid, E., & Walster, E. *Interpersonal attraction*. Reading, Mass.: Addison-Wesley, 1969.
- Blauner, R. The auto worker and the assembly line: Alienation intensified. In R. Blauner, *Alienation and freedom: The factory worker and his industry*. Chicago: University of Chicago Press, 1964.
- Brehm, J. *A theory of psychological reactance*. New York: Academic Press, 1966.
- Burnstein, E., & Worchel, P. Arbitrariness of frustration and its consequences for aggression in a social situation. *Journal of Personality*, 1962, 30, 328-340.
- Buss, A. *The psychology of aggression*. New York: Wiley, 1961.
- Buss, A. Physical aggression in relation to different frustrations. *Journal of Abnormal and Social Psychology*, 1963, 67, 1-7.
- Camus, A. *The rebel*. New York: Knopf, 1956.
- Cofer, C., & Appley, M. *Motivation: Theory and research*. New York: Wiley, 1964.
- Cohen, A. Social norms, arbitrariness of frustration, and status of the agent of frustration in the frustration-aggression hypothesis. *Journal of*

- Abnormal and Social Psychology*, 1955, 51, 222-226.
- Dollard, J., Doob, L., Miller, N., Mowrer, O., & Sears, R. *Frustration and aggression*. New Haven: Yale University Press, 1939.
- Durkheim, E. *Suicide*. New York: Free Press, 1951. (Originally published, 1897.)
- Etzioni, A. Basic human needs, alienation, and inauthenticity. *American Sociological Review*, 1968, 33, 870-885.
- Festinger, L. *A theory of cognitive dissonance*. Stanford, Calif.: Stanford University Press, 1957.
- Forward, J., & Williams, J. Internal-external control and black militancy. *Journal of Social Issues*, 1970, 26, 75-92.
- Groat, H., & Neal, A. Social psychological correlates of urban fertility. *American Sociological Review*, 1967, 32, 945-959.
- Gurin, G., & Gurin, P. Expectancy theory in the study of poverty. *Journal of Social Issues*, 1970, 26, 83-104.
- Heider, F. Attitudes and cognitive organization. *Journal of Psychology*, 1946, 21, 107-112.
- Heider, F. *The psychology of interpersonal relations*. New York: Wiley, 1958.
- Horwitz, M. The recall of interrupted group tasks: An experimental study of individual motivation in relation to group goals. *Human Relations*, 1954, 7, 3-38.
- Jones, E., & Gerard, H. *Foundations of social psychology*. New York: Wiley, 1967.
- Keniston, K. *The uncommitted: Alienated youth in American society*. New York: Dell, 1965.
- Kierkegaard, S. *Either/Or* (2 vols.). New York: Anchor Books, 1959.
- Lazarus, R. *Psychological stress and the coping process*. New York: McGraw-Hill, 1966.
- Levine, S., & Scotch, N. (Eds.), *Social stress*. Chicago: Aldine, 1970.
- Lewis, O. *The children of Sanchez*. New York: Random House, 1961.
- Marx, K. Economic and philosophical manuscripts. In T. B. Bottomore (Trans. & Ed.), *Early writings*. New York: McGraw-Hill, 1964. (Originally published, 1844.)
- Marx, K., & Engels, F. Manifesto of the Communist Party. In L. Feuer (Ed.), *Basic writings on politics and philosophy*. New York: Doubleday, 1959. (Originally published, 1848.)
- McGrath, J. *Social and psychological factors in stress*. New York: Holt, Rinehart & Winston, 1970.
- Merton, R. Social structure and anomie. *American Sociological Review*, 1938, 3, 672-682.
- Pastore, N. The role of arbitrariness in the frustration-aggression hypothesis. *Journal of Abnormal and Social Psychology*, 1952, 47, 728-731.
- Ransford, H. Isolation, powerlessness, and violence: A study of attitudes and participation in the Watts riot. *American Journal of Sociology*, 1968, 73, 581-591.
- Rothaus, P., & Worchel, P. The inhibition of aggression under nonarbitrary frustration. *Journal of Personality*, 1960, 28, 108-117.
- Sarbin, T. The culture of poverty, social identity, and cognitive outcomes. In V. Allen (Ed.), *Psychological factors in poverty*. Chicago: Markham, 1970.
- Sartre, J. P. *Being and nothingness*. New York: Washington Square Press, 1953.
- Seeman, M. On the meaning of alienation. *American Sociological Review*, 1959, 24, 783-791.
- Simmel, G. The metropolis and mental life. In G. Simmel, *The sociology of Georg Simmel*. Glencoe, Ill.: Free Press, 1950. (Originally published, 1902).
- Stokols, D. On the distinction between density and crowding: Some implications for future research. *Psychological Review*, 1972, 79, 275-277.
- Stokols, D. Some determinants of alienation in the small group. Unpublished doctoral dissertation, University of North Carolina, Chapel Hill, 1973.
- Thibaut, J., Friedland, N., & Walker, L. Compliance with rules: Some social determinants. *Journal of Personality and Social Psychology*, 1974, 30, 792-801.
- Thibaut, J., & Kelley, H. *The social psychology of groups*. New York: Wiley, 1959.
- Weber, M. Bureaucracy. In H. Gerth & C. W. Mills (Eds.), *From Max Weber*. New York: Oxford University Press, 1958.
- Wirth, L. Urbanism as a way of life. *American Journal of Sociology*, 1938, 44, 1-24.
- Worchel, S., & Brehm, J. Effect of threats to attitudinal freedom as a function of agreement with the communicator. *Journal of Personality and Social Psychology*, 1970, 14, 18-22.

(Received March 8, 1974)

Sentence Comprehension: A Psycholinguistic Processing Model of Verification

Patricia A. Carpenter and Marcel Adam Just
Carnegie-Mellon University

This paper presents a theoretical account of certain aspects of sentence comprehension, particularly the processes involved in deciding if a sentence is true or false. The model proposes that (a) sentences are internally represented as an ordered set of constituents in an abstract propositional format, (b) other information sources (like pictures) may be represented in a similar format in this task, and (c) the corresponding constituents from the sentence representation and the picture representation are serially compared. The model accounts for verification latencies for a class of complex sentences that differ in the scope of the negation operator (e.g., *It isn't true that the dots are red* versus *It's true that the dots aren't red*). The predictions are made on the basis of a single parameter—the time to find and compare one pair of constituents. The model also accounts for verification of counterfactual clauses, implicit negatives, universal and particular quantifiers, and sentence recoding. The formal structure of the model seems appropriate for dealing with a broad domain of language processing. Since the model is expressed in terms of basic mental operations and is incorporated into a more global theory of the immediate processor, it integrates sentence comprehension with the more general area of cognitive psychology.

A major challenge to cognitive science has been to explain language comprehension within the same information processing framework that characterizes other complex human behavior, such as memory scanning (cf. Sternberg, 1969a), concept formation (cf. Hunt, Marin, & Stone, 1966), perception (cf. Haber & Hershen-son, 1973), and problem solving (cf. Newell & Simon, 1972). This article builds on the same theoretical framework to explain the mental processes in sentence comprehension. The main focus will be on the

The order of authors is arbitrary. This paper represents a collaborative effort.

We would like to thank Bill Chase, Al Newell, and Herb Simon for their comments on earlier drafts of this article and Herb Clark for his comments and his intellectual support and interest in our work.

The project presented herein was partially supported by a grant from the National Institute of Education, U.S. Department of Health, Education, and Welfare, Grant NIE-G-74-0016. The research was also partially supported by the National Institute of Mental Health Grants MH-07722 and MH-20021.

Requests for reprints should be sent to Patricia A. Carpenter, Psychology Department, Carnegie-Mellon University, Pittsburgh, Pennsylvania 15213.

processes involved in verifying a sentence as true or false. The model deals with the internal representation of a sentence in short-term memory and the operations applied to that representation in order to verify it. The verification process will be examined in terms of elementary mental operations such as the comparison of two elements (Posner, 1969) and control processes that govern the sequence in which operations are performed (Newell 1972, 1973). The analytic methods to be used in constructing the model treat response latencies as measures of the duration of mental processes (Sternberg, 1969a, 1969b).

In the first section of the article, we will develop a model of the verification process and show how it accounts for the comprehension of simple affirmative and negative sentences. We will also provide a rigorous test of the model by examining the verification of embedded affirmative and negative sentences and by testing the model against a number of previous studies. The second section presents a related model of a different kind of sentence processing—recoding a sentence representation from one form into an equivalent form. The

recoding model is based on operations and parameters similar to those in the main model of verification. Finally, in the third section, we will evaluate the generality of the model by examining the comprehension of a variety of semantic constructions, such as quantifiers, counterfactuals, implicit negatives, and active and passive sentences.

The model is most easily explained by first outlining the kind of verification task and results that motivated its development. In a typical verification task (Just & Carpenter, 1971, Experiment 2), a person was first shown an array of dots, all of which were either red or black, and then a sentence like *The dots are red* or *The dots aren't red*. And, he was timed while he looked at the array, read the sentence, and decided if it was true or false of the picture. So on various trials, the sentence could be affirmative or negative and either true or false. The two main results of the study are easily summarized. First, there was an interaction between the kind of sentence (affirmative or negative) and the type of response. True affirmatives were verified faster than false affirmatives, but false negatives were verified faster than true negatives. The second main result was that negative sentences took longer to verify. Both of these results are common to many other studies (Chase & Clark, 1972; Clark & Chase, 1972; Gough, 1965, 1966; McMahon, 1963; Slobin, 1966; Trabasso, Rollins, & Shaughnessy, 1971; Wason & Jones, 1963).

To explain the interaction between affirmative-negative and true-false, Gough (1965) suggested that it is easier to compare the representation of the sentence to that of the picture when the two are congruent. For example, in the case of the affirmative sentences, the color predicate of the sentence matches the accompanying picture in the true case (e.g., *The dots are red* paired with a picture of red dots); however, the colors mismatch in the false case (e.g., *The dots are red* paired with black dots). Congruence in the true affirmative case explains why it is easier to process than the false affirmative case. For negative sentences, the match occurs

in the false case (e.g., *The dots aren't red* paired with red dots) and the mismatch occurs in the true case (e.g., *The dots aren't red* paired with black dots). Therefore, congruence also explains why false negatives are verified faster than true ones. In summary, the comparison between sentence and picture is facilitated if the corresponding elements match and takes longer if they mismatch. The extra time consumed by a color mismatch is called *falsification time*, and its magnitude is approximately the same for affirmatives and negatives.

Congruence relations between the sentence and picture representations will also explain the second main result, why negative sentences take longer to verify (Chase & Clark, 1972; Trabasso et al., 1971). Suppose the representation of a negative sentence, like *The dots aren't red*, contains a marker to denote the negative polarity of the predication. The negative marker will mismatch with the affirmative picture representation. This mismatch in the negative cases causes the comparison process to take longer than in affirmative cases, where both the sentence and picture are affirmative. The difference in latencies between negative and affirmative sentences is called *negation time*.

In summary, the two parameters, negation time and falsification time, account for the verification latencies. Both parameters may reflect extra processing time caused by mismatches between corresponding components of the sentence and picture representations. The amount of extra processing time was called falsification time or negation time, depending on whether the predicates or polarity markers mismatched.

We examined the data from previous experiments and discovered a very persistent relation between the values of these two parameters. Negation time was either two or four times greater than falsification time. The discovery that negation time is an integral multiple of falsification time suggested that both parameters are determined by a single mental operation, and what varies is the number of times the

operation is executed. This suggestion led to the following model for how the information in sentences is represented and compared to other information.

THE CONSTITUENT COMPARISON MODEL

The proposed model postulates that a sentence and picture are represented in abstract propositional formats. To verify the sentence, the constituents in the two representations are compared—tested for identity at an abstract level. The number of operations in the comparison process is the primary determinant of the verification latency.

The internal representation. Understanding a sentence involves internally representing the information that the sentence contains. It is likely that the format of the internal representation is propositional, a relational structure consisting of a predicate and one or more arguments (cf. Clark & Chase, 1972; Crothers, 1972; Kintsch, 1972; Rumelhart, Lindsay, & Norman, 1972; Simon, 1972). We will use the conventional notation, (PREDICATE, ARGUMENT), to denote a proposition. So, part of the representation of a simple declarative sentence like *The dots are red* would be (RED, DOTS), meaning that redness is predicated of the dots. Predications can be affirmed or negated. So, the entire representation of this affirmative sentence would be [AFF, (RED, DOTS)]. A negative sentence like *The dots aren't red* would be represented as [NEG, (RED, DOTS)]. AFF and NEG are embedding markers that denote the affirmative or negative polarity of the predication. This form of representation allows us to combine simple propositions to represent more complex sentences. For example, *It is fortunate that the dots are red* might be represented as {FORTUNATE, [AFF, (RED, DOTS)]}.

The internal representation of a sentence is not necessarily linguistic in nature. The verbal symbols in these representations, for example, DOTS, are used to denote more abstract entities. In fact, research on sentence-picture verification suggests

that there may be a level of representation that is neither linguistic nor pictorial in nature but can represent information from either domain (Chase & Clark, 1972; Clark & Chase, 1972). (But this does not imply that pictures cannot be represented otherwise in other situations, cf. Cooper and Shepard, 1973). In these verification tasks, a propositional representation may appropriately characterize the information encoded from a picture. For example, a picture of red dots may be represented (RED, DOTS). We assume that pictures are generally represented affirmatively, but by convention, we have deleted the affirmative polarity marker. The presence or absence of this marker in the picture representation has no consequences for any of our proposals, so long as it is assumed that the absence of a marker denotes affirmative polarity. The picture representation may be influenced by the sentence that precedes it. In particular, if the picture has several possible interpretations, it is coded so that it has the same argument as the sentence representation (Carpenter & Just, 1972; Carpenter & Just, in press; Just & Carpenter, 1971).

The detailed form of the representation of various kinds of sentences has not yet been empirically verified. For example, the research on the linguistic factors that determine the psychological predicate-argument structure has only begun (cf. Halliday, 1967; Hornby, 1972). Moreover, there are cases where various representations are formally equivalent, and the selection of one particular form is somewhat arbitrary. Nevertheless, our conventional notation is sufficient for the current model and promises to be flexible enough to accommodate a variety of linguistic structures (cf. Kintsch, 1972).

The same sentence may be represented differently in different situations. This follows from the assumption that the representation contains the information that a person extracts from a sentence. What information is extracted depends on the preceding sentences, the situation in which the sentence is embedded, and the listener's

recoding model is based on operations and parameters similar to those in the main model of verification. Finally, in the third section, we will evaluate the generality of the model by examining the comprehension of a variety of semantic constructions, such as quantifiers, counterfactuals, implicit negatives, and active and passive sentences.

The model is most easily explained by first outlining the kind of verification task and results that motivated its development. In a typical verification task (Just & Carpenter, 1971, Experiment 2), a person was first shown an array of dots, all of which were either red or black, and then a sentence like *The dots are red* or *The dots aren't red*. And, he was timed while he looked at the array, read the sentence, and decided if it was true or false of the picture. So on various trials, the sentence could be affirmative or negative and either true or false. The two main results of the study are easily summarized. First, there was an interaction between the kind of sentence (affirmative or negative) and the type of response. True affirmatives were verified faster than false affirmatives, but false negatives were verified faster than true negatives. The second main result was that negative sentences took longer to verify. Both of these results are common to many other studies (Chase & Clark, 1972; Clark & Chase, 1972; Gough, 1965, 1966; McMahon, 1963; Slobin, 1966; Trabasso, Rollins, & Shaughnessy, 1971; Wason & Jones, 1963).

To explain the interaction between affirmative-negative and true-false, Gough (1965) suggested that it is easier to compare the representation of the sentence to that of the picture when the two are congruent. For example, in the case of the affirmative sentences, the color predicate of the sentence matches the accompanying picture in the true case (e.g., *The dots are red* paired with a picture of red dots); however, the colors mismatch in the false case (e.g., *The dots are red* paired with black dots). Congruence in the true affirmative case explains why it is easier to process than the false affirmative case. For negative sentences, the match occurs

in the false case (e.g., *The dots aren't red* paired with red dots) and the mismatch occurs in the true case (e.g., *The dots aren't red* paired with black dots). Therefore, congruence also explains why false negatives are verified faster than true ones. In summary, the comparison between sentence and picture is facilitated if the corresponding elements match and takes longer if they mismatch. The extra time consumed by a color mismatch is called *falsification time*, and its magnitude is approximately the same for affirmatives and negatives.

Congruence relations between the sentence and picture representations will also explain the second main result, why negative sentences take longer to verify (Chase & Clark, 1972; Trabasso et al., 1971). Suppose the representation of a negative sentence, like *The dots aren't red*, contains a marker to denote the negative polarity of the predication. The negative marker will mismatch with the affirmative picture representation. This mismatch in the negative cases causes the comparison process to take longer than in affirmative cases, where both the sentence and picture are affirmative. The difference in latencies between negative and affirmative sentences is called *negation time*.

In summary, the two parameters, negation time and falsification time, account for the verification latencies. Both parameters may reflect extra processing time caused by mismatches between corresponding components of the sentence and picture representations. The amount of extra processing time was called falsification time or negation time, depending on whether the predicates or polarity markers mismatched.

We examined the data from previous experiments and discovered a very persistent relation between the values of these two parameters. Negation time was either two or four times greater than falsification time. The discovery that negation time is an integral multiple of falsification time suggested that both parameters are determined by a single mental operation, and what varies is the number of times the

operation is executed. This suggestion led to the following model for how the information in sentences is represented and compared to other information.

THE CONSTITUENT COMPARISON MODEL

The proposed model postulates that a sentence and picture are represented in abstract propositional formats. To verify the sentence, the constituents in the two representations are compared—tested for identity at an abstract level. The number of operations in the comparison process is the primary determinant of the verification latency.

The internal representation. Understanding a sentence involves internally representing the information that the sentence contains. It is likely that the format of the internal representation is propositional, a relational structure consisting of a predicate and one or more arguments (cf. Clark & Chase, 1972; Crothers, 1972; Kintsch, 1972; Rumelhart, Lindsay, & Norman, 1972; Simon, 1972). We will use the conventional notation, (PREDICATE, ARGUMENT), to denote a proposition. So, part of the representation of a simple declarative sentence like *The dots are red* would be (RED, DOTS), meaning that redness is predicated of the dots. Predications can be affirmed or negated. So, the entire representation of this affirmative sentence would be [AFF, (RED, DOTS)]. A negative sentence like *The dots aren't red* would be represented as [NEG, (RED, DOTS)]. AFF and NEG are embedding markers that denote the affirmative or negative polarity of the predication. This form of representation allows us to combine simple propositions to represent more complex sentences. For example, *It is fortunate that the dots are red* might be represented as {FORTUNATE, [AFF, (RED, DOTS)]}.

The internal representation of a sentence is not necessarily linguistic in nature. The verbal symbols in these representations, for example, DOTS, are used to denote more abstract entities. In fact, research on sentence-picture verification suggests

that there may be a level of representation that is neither linguistic nor pictorial in nature but can represent information from either domain (Chase & Clark, 1972; Clark & Chase, 1972). (But this does not imply that pictures cannot be represented otherwise in other situations, cf. Cooper and Shepard, 1973). In these verification tasks, a propositional representation may appropriately characterize the information encoded from a picture. For example, a picture of red dots may be represented (RED, DOTS). We assume that pictures are generally represented affirmatively, but by convention, we have deleted the affirmative polarity marker. The presence or absence of this marker in the picture representation has no consequences for any of our proposals, so long as it is assumed that the absence of a marker denotes affirmative polarity. The picture representation may be influenced by the sentence that precedes it. In particular, if the picture has several possible interpretations, it is coded so that it has the same argument as the sentence representation (Carpenter & Just, 1972; Carpenter & Just, in press; Just & Carpenter, 1971).

The detailed form of the representation of various kinds of sentences has not yet been empirically verified. For example, the research on the linguistic factors that determine the psychological predicate-argument structure has only begun (cf. Halliday, 1967; Hornby, 1972). Moreover, there are cases where various representations are formally equivalent, and the selection of one particular form is somewhat arbitrary. Nevertheless, our conventional notation is sufficient for the current model and promises to be flexible enough to accommodate a variety of linguistic structures (cf. Kintsch, 1972).

The same sentence may be represented differently in different situations. This follows from the assumption that the representation contains the information that a person extracts from a sentence. What information is extracted depends on the preceding sentences, the situation in which the sentence is embedded, and the listener's

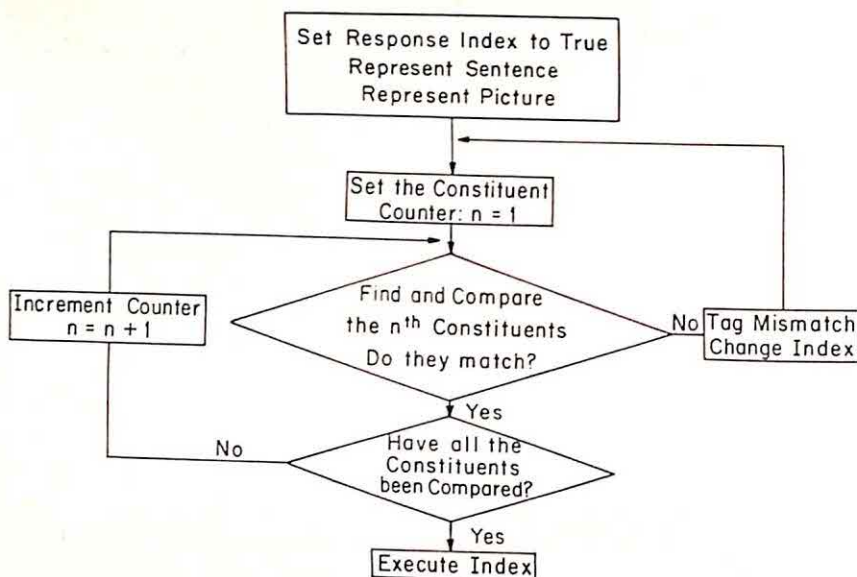


FIGURE 1. The constituent comparison model.

previous knowledge. In other words, context plays a role in how a sentence is represented, suggesting an important distinction between the psychological notion of an internal representation and the traditional linguistic notion of deep structure.

The mental operations. The main focus of the model is on the operations that compare the sentence and picture representations. The model postulates that the corresponding constituents from the two representations are retrieved and compared, pair by pair. Moreover, the number of these retrieve-and-compare operations is assumed to be the primary determinant of the pattern of verification latencies. Figure 1 shows the proposed process in flowchart form. The representations' propositional structure and embeddings provide an ordering relation on the constituents. This ordering determines the sequence in which constituents are compared. Inner propositions are compared before polarity markers. An AFF marker in a sentence representation is assumed to match the absence of a marker in the picture representation, since pictures are generally encoded affirmatively. The "find-and-compare" process is a serial, iterative operation that can be applied to representations with multiple embeddings.

This iterative operation will allow the model to be generalized without additional assumptions.

A central assumption is that whenever two corresponding constituents from the sentence and picture representation mismatch, then the entire comparison process is reinitialized. To prevent the process from looping forever on mismatching constituents, we assume that the first time a mismatch is discovered, the two constituents involved are tagged, so that on subsequent recomparisons the two will be treated as a match.

Since mismatches cause the comparison process to be reinitialized, the total number of comparison operations, and consequently the total latency, increases with the number of mismatches. Moreover, a mismatch that occurs later in the comparison process results in more recomparisons than a mismatch on earlier constituents. So the total latency is a function of both the number of mismatches and their locus in their respective representations.

A response index monitors the matches and mismatches between constituents. The index has two possible states, *true* and *false*. At the beginning of each trial, its initial state is *true*, but each mismatch causes it to change from its current state to its other state. The time spent in

changing the response index (and for that matter, tagging mismatching constituents) is assumed to be negligible relative to the time to perform the find-and-compare operation.

When the model is applied to the proposed sentence and picture representations in the Just and Carpenter (1971) experiment, it can account for the latencies in the four conditions, as shown in Table 1. In Table 1, a match between two constituents is designated by a plus (+) under them, and a mismatch by a minus (-). In this notation, the comparisons proceed from right to left until a mismatch is encountered. A reinitialization is denoted by a new line of pluses and minuses. In the simplest case, the true affirmative, there are no mismatches between the sentence and picture representations. The first comparison, between the inner propositions, results in a match. The second comparison, between polarity markers, also results in a match. (The AFF marker in the sentence representation is presumed to match the absence of any polarity marker in the picture representation). Thus after a total of two constituent com-

parisons, the truth index is still set to *true*, and this response is executed. The number of constituent comparisons in the true affirmative case serves as the baseline for the other conditions and will be referred to as k . Here, k equals 2.

In the false affirmative condition, the inner propositions of the sentence and picture mismatch. The mismatch will re-initialize the comparison process, causing one extra comparison above the base number. Table 2 shows the consequences of this mismatch in detail. The mismatching constituents are marked with a tag (denoted in Table 2 with an M), and the response index is set to *false*. After reinitialization, the tagged inner constituents are compared, and they match. The next comparison, between the polarity markers, also results in a match. So, the response *false* is executed after a total of $k + 1$ constituent comparisons.

In the false negative condition, there will be a total of $k + 2$ comparisons, due to the mismatch on the second constituent, the polarity markers. This mismatch will cause the response *false* to be executed. For the true negative condition, both the

TABLE 1
REPRESENTATIONS AND PREDICTIONS FOR THE FOUR INFORMATION CONDITIONS

Stimulus and representation	True affirmative	False affirmative
Sentence Picture Sentence representation Picture representation	The dots are red. Red dots [AFF, (RED, DOTS)] (RED, DOTS) + + response = true k comparisons	The dots are red. Black dots [AFF, (RED, DOTS)] (BLACK, DOTS) - + + response = false $k + 1$ comparisons index = false
	False negative	True negative
Sentence Picture Sentence representation Picture representation	The dots aren't red. Red dots [NEG, (RED, DOTS)] (RED, DOTS) - + + + response = false $k + 2$ comparisons index = false	The dots aren't red. Black dots [NEG, (RED, DOTS)] (BLACK, DOTS) - - + + + response = true $k + 3$ comparisons index = false index = true

Note. Plus and minus signs denote matches and mismatches of the corresponding constituents. Each horizontal line of plus and minus signs indicates a reinitialization of the comparison process.

TABLE 2

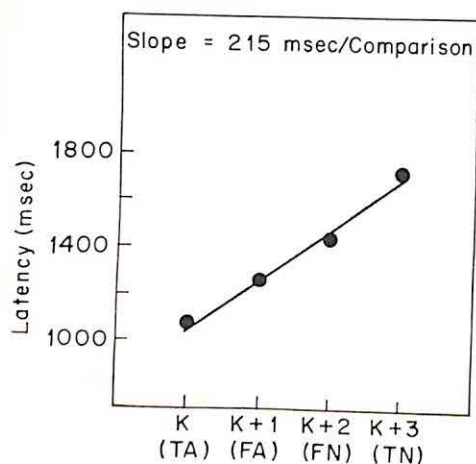
A TRACE OF THE OPERATIONS IN VERIFYING A FALSE AFFIRMATIVE

Stimulus	
Sentence	<i>The dots are red.</i>
Picture	A set of black dots
Operations	Trace
Initialize response index to <i>true</i>	
Represent sentence	[AFF, (RED, DOTS)]
Represent picture	(BLACK, DOTS)
Compare first constituents	—
Tag sentence constituent	[AFF, (M)]
Tag picture constituent	(M)
Change index to <i>false</i>	
Reinitialize comparison process	
Compare first constituents	+
Compare second constituents	+
Respond with content of index	<i>False</i>
Number of comparisons	$k + 1$, where $k = 2$

Note. Abbreviation: M denotes a neutral tag.

first and the second constituents mismatch so the response *true* will be executed after a total of $k + 3$ constituent comparisons.

The model postulates that verification latencies should be a direct function of the number of constituent comparisons.



NUMBER OF CONSTITUENT COMPARISONS

FIGURE 2. The fit of the constituent comparison model to the data from Just and Carpenter (1971). (Abbreviations: TA = true affirmative, FA = false affirmative, FN = false negative, TN = true negative.)

The number of comparisons, and hence the latency, should increase linearly from true affirmatives (k) to false affirmatives ($k + 1$) to false negatives ($k + 2$) to true negatives ($k + 3$).

The verification latencies for *The dots are(n't) red*, as well as the best fitting straight line, are shown in Figure 2 (data from Just & Carpenter, 1971, Experiment 2). The predictions of the model fit the data quite well. The model accounts for 98.0% of the variance among the four means. The slope is 215 msec per constituent comparison.

Only one parameter, the time to find and compare a pair of constituents, is necessary to describe the mental operations that are involved in these four conditions.¹ Thus, the current model is a very parsimonious explanation of these latencies. Although parsimony is desirable, the model should be evaluated on the basis of its ability to offer a rigorous formulation that both is a plausible mental process and can incorporate a wide variety of empirical results. This ability will be demonstrated in each of the following sections. The nature of the propositional representation, particularly the embedding feature, should allow the model to be generalized to more complex sentences without additional modifications. This property of the model was tested by examining embedded affirmative and negative sentences.

Tests of the Model

Experiment 1

We devised a more stringent test of the model by systematically varying the scope of the negative—the range of constituents to which the negative applies (Jackendoff,

¹ Of course, a second parameter for the intercept is necessary to describe a straight line. The intercept time is a wastebasket parameter, containing the durations of all of the processes which are common to all the conditions. In this spirit, we will refer to the "one" parameter of our model, namely the slope, when there are really two. Similarly, we will refer to the "two" parameters of the Clark and Chase (1972) and Trabasso et al. (1971) model when there are really three.

1969; Klima, 1964; Weinreich, 1963). We varied the scope by using affirmative sentences with the superordinate clause, *It is true that . . .*, for example, *It is true that the dots are red*. Such an affirmative can be negated in two ways. With one type of negation, the negative has a small scope, namely the inner predication: *It is true that the dots aren't red*. This will be called predicate negation. The second type of negation has a large scope since the negative is in the superordinate clause where it negates the entire inner proposition: *It isn't true that the dots are red*. This type of negation will be called denial. These two kinds of negation may be represented differently, and they should have different verification latencies. The precise predictions can be derived by examining how the comparison process might operate on the hypothesized representations for these types of sentences.²

The representation of an affirmative sentence like *It's true that the dots are red* may be the same as for the simple sentence *The dots are red*, namely, [AFF, (RED, DOTS)]. The rationale is that the embedding clause *It's true . . .* does not change the truth value and so it can be ignored. To demonstrate this point, consider a concatenation of this type of clause, for example, *It's true that it's true that it's true . . . the dots are red*. The number of such embedding clauses is irrelevant to the truth value of the proposition. Similarly, the embedding affirmative clause may be deleted from the representation of a predicate negative sentence like *It's true that the dots aren't red*, so that the representation would be [NEG, (RED, DOTS)]. However, the representation of a denial like *It isn't true that the dots are red* must include the negative embedding clause. Here the embedding clause does affect the truth value of the sentence. Thus, denial sentences might be represented {NEG, [AFF, (RED, DOTS)]}. The pictures would be represented as simple propositions like (RED, DOTS) or (BLACK, DOTS). Table 3

shows examples of the representations in the six conditions.

The predictions of the model can be derived by examining the flowchart in Figure 1 and the representations in Table 3. Table 3 details the process, showing the comparisons and outcomes for each of the six information conditions. The predicted number of operations necessary to verify a true affirmative is k ; a false affirmative is $k + 1$; a false predicate negative is $k + 2$; a true predicate negative is $k + 3$; a false denial is $k + 4$; and a true denial is $k + 5$. The verification latencies should increase linearly with the proposed number of operations. If this result is obtained, it will constitute strong support for the constituent comparison model and the notion of a single underlying iterative operation.

Method. The experiment was a verification task in which the subject was timed while he read a sentence, looked at a picture, and then decided whether the sentence was true or false of the picture.

The stimulus sentences were the three in Table 3, as well as six more in which the adjectives *black* or *green* replaced the adjective *red*. The picture was an array of 16 dots of one color, either red, green, or black. Table 3 shows the truth values for the various combinations of sentences and pictures. Each affirmative sentence could be false with respect to two pictures and true with respect to one picture; consequently, the one true sentence-picture combination was presented twice. The sentence-picture combinations for the false negative conditions were presented twice for the same reason. Thus, there was a total of 36 sentence-picture combinations. Each sentence was typed in elite type and subtended 10° of visual angle. The 4×4 array of dots, which was drawn below the sentence with a felt tip pen, subtended $2.5^\circ \times 2.5^\circ$ of visual angle. The stimulus card was viewed in a tachistoscope at a distance of 58.4 cm.

The subject initiated a trial by pushing a switch, and 500 msec later the stimulus card was presented and remained in view until the subject responded. The subject's decision was made using a two-button decision apparatus. The assignment of dominant hand to *true* button was balanced across subjects. The subject was given feedback about the correctness of his answer only during the practice session which consisted of 10 trials selected at random from the ensemble of 36. After the practice, all 36 cards were shuffled and the subject went through 7 blocks of 36 trials. The testing

² These two types of negatives have the same truth value here, although this is not always the case (Fillmore, 1963; Lakoff, 1970).

TABLE 3
REPRESENTATIONS AND PREDICTIONS FOR THE SIX INFORMATION CONDITIONS

Stimuli and representations	True affirmative	False affirmative
Sentence Picture Sentence representation Picture representation	It's true that the dots are red. Red dots [AFF, (RED, DOTS)] (RED, DOTS) + + response = true k comparisons	It's true that the dots are red. Black dots [AFF, (RED, DOTS)] (BLACK, DOTS) - index = false + + response = false $k + 1$ comparisons
	False predicate negative	True predicate negative
Sentence Picture Sentence representation Picture representation	It's true that the dots aren't red. Red dots (NEG, (RED, DOTS)] (RED, DOTS) - + index = false + + response = false $k + 2$ comparisons	It's true that the dots aren't red. Black dots [NEG, (RED, DOTS)] (BLACK, DOTS) - index = false + index = true + + response = true $k + 3$ comparisons
	False denial	True denial
Sentence Picture Sentence representation Picture representation	It isn't true that the dots are red. Red dots {NEG, [AFF, (RED, DOTS)]} (RED, DOTS) - + + index = false + + + response = false $k + 4$ comparisons	It isn't true that the dots are red. Black dots {NEG, [AFF, (RED, DOTS)]} (BLACK, DOTS) - index = false + index = true + + + response = true $k + 5$ comparisons

Note. Plus and minus signs denote matches and mismatches of the corresponding constituents. Each horizontal line of plus and minus signs indicates a reinitialization of the comparison process.

session lasted about 50 minutes. The subjects were 24 students of Carnegie-Mellon University who were paid for their services.

Results. The latencies for erroneous responses (3.2%) were discarded, and the subject's score for each condition was the mean of his correct latencies for that condition. As predicted, the mean latencies increased linearly with the number of hypothesized constituent comparisons. More precisely, latencies increased an average of 200 msec for each additional constituent comparison (standard error = 23 msec). Figure 3 shows this result, along with the best fitting straight line. The model accounts for 97.7% of the vari-

ance among the six means, $F(1, 115) = 171.17, p < .01$. The residual 2.3% is not significant, $F(4, 115) = 1.01$. The root mean squared deviation of 52 msec is small relative to the 200 msec parameter. This analysis confirms the major hypothesis that verification time increases linearly with the number of constituent comparisons.

The error rates for the six conditions were correlated with the latencies ($r = .98$), as is shown in Figure 3. This correlation indicates that the probability of error increases with the number of hypothesized operations.

Discussion. The model predicts the processing time for these six conditions on

the basis of a single parameter: the time to find and compare a pair of constituents. This strongly supports the hypothesis that a single iterative operation underlies the processing of affirmative and negative sentences. These data also support the hypothesis that the two negatives are represented differently and that the essence of this difference lies in the locus of the negative in the internal representation.

A further control study showed that the representation and processing of the sentence is determined by its semantics rather than by its surface structure. To show this, we compared the processing of sentences that had the same constituent structure but different surface structures. In this control study, the inner propositions of the sentences were embedded in two ways: the same way as in Experiment 1 (e.g., *It's true that the dots aren't red*) and with the embedding clause at the end of the sentence (e.g., *That the dots aren't red is true*). Both of these kinds of sentences are postulated to have the same underlying constituent structure. However, the position of the negative in the surface structure has been changed. If the results of the basic experiment can be replicated with the new sentences, then the results must be due to the postulated representation rather than the surface structure. The new stimulus sentences were: Affirmative—*That the dots are red is true*; predicate negative—*That the dots aren't red is true*; and denial—*That the dots are red isn't true*. Thus, there were three basic sentence types, two kinds of surface structures and the true-false variable. In addition, each sentence had the adjective either *red* or *black*, for a total of 24 stimuli. The pictures were 4 × 4 arrays of all red or all black dots. The 12 paid subjects were run for 2 blocks of 24 trials with a procedure identical to that of Experiment 1.

This control study showed that the two types of surface structures were processed similarly. Both types of sentences showed a linear increase in latency as the number of comparisons increased. Regardless of whether the negative morpheme occurred

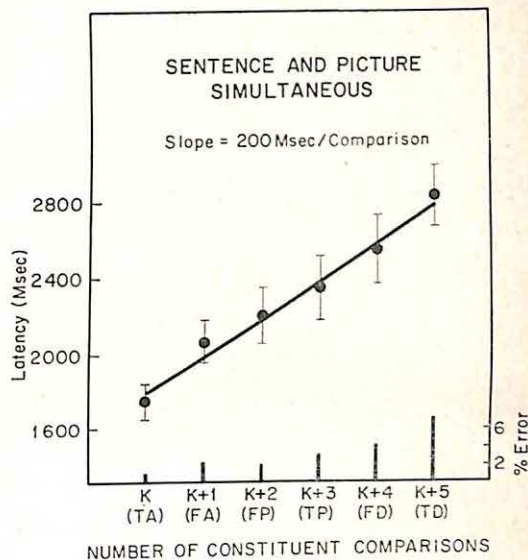


FIGURE 3. The fit of the constituent comparison model to the results of Experiment 1. (Abbreviations: TA = true affirmative, FA = false affirmative, FP = false predicate negative, TP = true predicate negative, FD = false denial, TD = true denial.)

near the beginning or end of the sentence, denials took about 500 msec longer to verify than predicate negatives. This result shows that the underlying constituent structure, rather than order of negatives in the surface structure, determines processing time and constitutes further support for the proposed representations for the two kinds of negative sentences.

The constituent comparison model correctly predicts that the pattern of latencies in this control experiment should be the same as in Experiment 1. Whether the embedding string was at the beginning or end of the sentence, there was a linear increase in latency with the number of hypothesized constituent comparisons. The best fitting straight line for sentences that were the same as in Experiment 1 had a slope of 234 msec/comparison, standard error = 28 msec. The linear trend was significant, $F(1, 121) = 55.24$, $p < .01$. The model accounts for 97.4% of the variance among the six means. The residual was not significant, $F(4, 121) < 1$. The best fitting line for the sentences with the new serial order had a slope of 286 msec/comparison, standard error = 40

msec. The linear trend was significant, $F(1, 121) = 82.77$, $p < .01$. The model accounts for 95.5% of the variance among the six means. The residual for this new serial order was also not significant, $F(4, 121) < 1$. Finally, the two slopes were not significantly different from each other, $F(1, 121) = 1.38$. Thus, the serial position of the negative in the surface structure of the sentence does not alter the underlying linearity of the results.

This control study also rules out an alternative interpretation of Experiment 1. In Experiment 1, the scope of the negative morpheme is correlated with its serial position in the surface structure of the sentence. The negative in a denial sentence occurs near the beginning of the sentence, while the negative in a predicate negative occurs at the end. Thus, the increase in verification latencies was correlated with the serial position of the negative morpheme. One might argue that negatives at the beginning of a sentence impose a greater memory load, and this greater memory load slows down subsequent comparison processes. This argument is countered by the result that the serial position of the negative is not responsible for the difference in negation time.

Experiment 2

Because of its focus on comparison processes, the model ignores the possibility that some sentences take longer to read than others. For example, it may take longer to read and represent a negative sentence than an affirmative one. Both the reading and representation time, as well as the comparison time, were included in the verification latencies in Experiment 1. Experiment 2 was designed to test whether latencies still increased linearly when reading and representation time were eliminated from the verification latencies. Subjects were given 2 sec to read and represent each sentence before the picture appeared. Verification time was measured only from the onset of the picture. If all reading and representation processes are complete after 2 sec, then any remaining differences in the sub-

sequent verification time can safely be attributed to the comparison process. It was decided that 2 sec was sufficient reading time on the basis of a small study where two naive subjects were timed as they read a sentence and indicated they had apprehended it by closing a microswitch.

Method. The materials and procedure were the same as in Experiment 1, except for the following difference. After the subject pushed the switch to initiate a trial, the sentence appeared in the upper channel of the tachistoscope for 2 sec and then disappeared as the picture appeared in the bottom channel. The subject was timed from the onset of the picture until he made a response. Thus, the response latencies should not include the time to read and represent the sentence. The subjects were 24 students at Carnegie-Mellon University who were paid for their services.

Results and discussion. The results were first analyzed for only the first block of trials, since previous pilot studies of such sentence-first conditions indicated a systematic change in the pattern of latencies across blocks of trials. The results for the later blocks and the justification for the partitioning of the data will be discussed in detail below.

The subject's score for each condition was the mean of his correct latencies for that condition. As predicted, the mean latencies increased linearly with the number of hypothesized constituent comparisons. Figure 4 shows this result, along with the best fitting straight line. The verification time increases an average of 169 msec for each additional constituent comparison (standard error = 21 msec). The model accounts for 97.9% of the variance among the six means, $F(1, 115) = 136.54$, $p < .01$. The residual 2.1% is not significant, $F(4, 115) < 1$. The root mean squared deviation of 42 msec is small relative to the 169 msec parameter. The error rate was low (6.5%). As in Experiment 1, the error rates for the six conditions were correlated ($r = .81$) with the latencies, as shown in Figure 4.

The data suggest that reading and representation time is a negligible component of the estimated time per operation in Experiment 1. The estimate there was 200 msec, 31 msec more than in Experi-

ment 2, where reading and representation time were eliminated. A t test comparing the slopes of the subjects in the two experiments showed no significant difference, $t(46) < 1$. Since the 200 msec slope involves seven times as much practice as the 169 msec slope, this comparison should be viewed with caution. In any case, it is clear that verification time increases linearly with the number of hypothesized constituent comparisons, regardless of whether reading and representation time is included.

The Control Structure of the Comparison Process

The processes that operate on the sentence and picture representations must have some control structure that governs the flow of activity in short-term memory (Atkinson & Shiffrin, 1968; Newell, 1972, 1973). We have written a computer simulation of the constituent comparison model and integrated it with a more global theory of the control structure of short-term memory, developed as a production system by Newell (1973). One purpose of the simulation was to make the control structure of the verification process explicit. A production system specifies exactly what conditions of short-term memory will trigger a particular operation and how that operation will affect the content of the memory.

A production system has two main components: (a) a short-term memory consisting of n ordered items, and (b) a list of independent production rules that operate on the items. Each rule consists of conditions and actions. Whenever the items in short-term memory satisfy the conditions of a particular rule, then the actions associated with that rule are performed. An action may change the contents of short-term memory by rearranging, adding, deleting, or modifying individual items. When the memory is filled to capacity, insertion of a new item causes displacement or forgetting of the oldest item. After an action is taken, the contents of the memory may satisfy the same rule or a different one. One rule

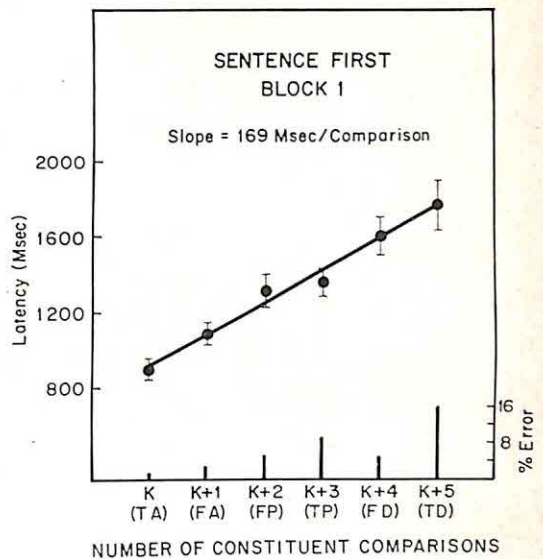


FIGURE 4. The fit of the constituent comparison model to the results of the first block of Experiment 2. (Abbreviations: TA = true affirmative, FA = false affirmative, FP = false predicate negative, TP = true predicate negative, FD = false denial, TD = true denial.)

after another is executed until no rule is satisfied or until a "stop" rule is executed.

At the beginning of the verification process, there are three items in the memory: a response index that is set to "true," a sentence representation, and a picture representation. Since these three items are essential for performing the task, they are repeatedly rehearsed, preventing them from being forgotten. There are five other items that play a role in the verification process. First of all, when two corresponding constituents are to be compared, a copy of each is retrieved from the sentence and picture representations. These copies constitute two additional items in the memory. Next, there is a place holder for each of the two representations that keeps track of which items were last retrieved and compared. Finally, another item, called "mismatch," is generated when two corresponding constituents mismatch. Thus, there are eight possible items involved in performing the task, and the memory is postulated to have a capacity of seven items.

Although the flowchart model and the production system both account for per-

formance in the same task, there are two respects in which the models differ. One difference concerns the reason for reinitializing the comparison process after a mismatch. The flowchart explanation treats the comparison process as one subroutine and tagging and changing the response index as a separate subroutine. Once the comparison subroutine is interrupted and control is transferred to the tagging subroutine, the comparison subroutine must start from the beginning. In the production system, the reinitialization results from the limitations of short-term memory. In short, there is a rule whose condition is a mismatch between two corresponding constituents. When the condition is satisfied, the action associated with the rule is performed, namely, the insertion of a "mismatch" item into memory. The presence of this "mismatch" item is the condition of a rule that updates the value of the response index. The "mismatch" item displaces a place holder that indicates which constituents were last compared. The absence of the place holder is eventually the condition of a rule that says, roughly, if there is no place holder, start the comparison process with the first constituent. This latter rule serves two purposes. It initializes the verification process at the beginning of a trial, and it reinitializes the process after a mismatch has displaced a place holder. In summary, a mismatch generates intervening items and operations, like updating the response index and tagging constituents. This makes the comparison process "forget" which constituents were last compared, and it consequently restarts from the beginning.

A second difference between the two models is in the nature of the "find-and-compare" operation. The flowchart model characterizes it as a unitary process. By contrast, the production system reveals that there are a number of suboperations in the find-and-compare operation, most of which can be classified as "find", and only one of which is "compare." This distinction between "find" and "compare" may be clarified by considering how we

mentally multiply two three-digit numbers (e.g., $245 \times 314 = ?$). In this task, most of the processing time seems to be spent in retrieving digits from various positions and keeping track of them; relatively little time seems to be spent in actually multiplying or adding pairs of digits (Danseman, 1969). Similarly, most of the processing time in sentence verification may be consumed by retrieving and positioning the appropriate constituents and keeping track of one's place; relatively little time may be spent in actually comparing items. If retrieval is a primary determinant of the total latency, this would explain why the time per find-and-compare operation is the same for all constituents. It is plausible that the time to retrieve a constituent may be independent of its type or complexity.

This retrieval plus comparison operation resembles the "scan-and-compare" operation in Sternberg's (1969a) short-term memory scanning model for the context recall task. At present, there is no way of empirically partitioning the slope into retrieval and comparison components either in memory-scanning tasks or in the sentence comprehension task.

A Test of the Model Against Previous Research

Since Wason's (1959) first work on verification, a fair number of studies have examined latencies in processing affirmative and negative sentences. These studies can now be reanalyzed in terms of the constituent comparison model. But before the model can be applied, the rules for representing negative sentences must be elaborated. Since these previous studies did not include explicit embedding strings like *It isn't true . . .*, negatives could have been represented with either a small scope of negation (as predicate negatives) or a large scope (as denials) (cf. Weinreich, 1963, p. 127). We have used the following decision rule, derived directly from the model, to determine the representation: If the ratio of negation time to falsification time was 2:1, then the negatives were represented with a small scope; but if the

ratio was 4:1, then the negatives were represented with a large scope. The rule can be derived by considering the relation of falsification time to negation time for the two scopes. Falsification time, due to a mismatch on the inner proposition, results from one additional constituent comparison. Negation time, with a small scope, is the result of two additional constituent comparisons, so the ratio of negation time to falsification time is 2:1. Negation time, with a large scope, is the result of four additional constituent comparisons, so the ratio is 4:1, and hence, the decision rule for previous studies.

When these representation rules were applied to previous verification studies, a surprising regularity emerged. In most studies where the sentence preceded the picture, the ratio of negation time to falsification time was 4:1, indicating that the negative was represented with a large scope. By contrast, in studies where the sentence followed the picture, the ratio was 2:1, indicating that the negative had a small scope. This regularity was particularly convincing in two studies that systematically examined the effect of reversing the sentence-picture order. In the Clark and Chase (1972) experiments, subjects were given sentences like *The star isn't above the plus* and pictures like (+) or (†). In the sentence-first condition, the ratio was 3.8:1, while it was 2.1:1 in the picture-first condition.³ In the second pair of experiments, Just and Carpenter (1971) presented a sentence like *The dots aren't red* followed or preceded by a picture of red dots or black dots. In the sentence-first condition the ratio was 4.2:1, while in the picture-first condition it was 1.7:1. These studies indicate that the same negative sentences were represented and processed with a different scope of negation, depending on whether they followed or preceded the picture.

³ We have analyzed the cases where the instruction for how to code the picture, "code what's on the top" or "code what's on the bottom," agreed with the sentence preposition (*above* and *below*, respectively). In some cases when they do not agree, there is recoding, an operation to be discussed later.

When the model's predictions and the actual data are laid out side by side, the fit is very convincing. As Table 4 shows, the model accounts very well for the comprehension of affirmative and negative sentences in a large number of previous studies.

The time to retrieve and compare a pair of constituents seems consistent within any one set of materials and experimental conditions, such as our own, or Clark and Chase's (1972) or Gough's (1965, 1966). However, there is some variability across various experiments and conditions. These data from the various experiments are based on different numbers of observations, and different sentences, pictures, etc.; so direct comparisons are difficult to make. However, it is plausible that the complexity of the linguistic or perceptual stimuli may affect the duration of a comparison. One factor that clearly does affect this duration is practice.

Practice effects. Extended practice in the sentence-verification task naturally makes performance much faster (Singer, Chase, Young, & Clark, Note 1). What is much more important, from our point of view, is that even though the time per operation decreases with practice, the linearity among the various information conditions is still preserved. In the Singer et al. experiment, three patient subjects verified sentences like *The star isn't above* for 15 days! The sentence preceded the picture, and timing started only from the onset of the picture. However, the pictures were of various types to discourage any sentence recoding prior to the picture. For example, the sentence *The star isn't above* could be true of the picture like + or one like †, where no star was present. Since these sentences are similar to those used by Clark and Chase (1972), and since it was a sentence-first experiment, it was assumed that the negative was represented with a large scope. The model predicts true affirmatives should require k operations, false affirmatives, $k + 1$; false negatives, $k + 4$; and true negatives, $k + 5$. In each 5-day period, latencies increased linearly with the number of constituent

TABLE 4
SUMMARY OF THE PREDICTIONS AND FIT OF THE CONSTITUENT COMPARISON MODEL

Reference	Example of a stimulus sentence	Percentage of variance accounted for by the model	Time per comparison (msec)	Observed (and estimated) latency						
Current sentence-first experiments										
Experiment 1				TA (k) ^a	FA (k + 1)	FP (k + 2)	TP (k + 3)	FD (k + 4)	TD (k + 5)	
Control	It is(n't) true that the dots are(n't) red.	97.7	200	1,737 (1,789)	2,073 (1,989)	2,211 (2,189)	2,344 (2,389)	2,538 (2,589)	2,830 (2,789)	
Control	It is(n't) true that the dots are(n't) red.	97.4	234	2,313 (2,209)	2,339 (2,443)	2,639 (2,677)	2,928 (2,911)	3,121 (3,145)	3,421 (3,379)	
Experiment 2, (Block 1)	That the dots are(n't) red is(n't) true.	95.5	286	1,972 (2,063)	2,325 (2,349)	2,863 (2,635)	2,860 (2,921)	3,197 (3,207)	3,452 (3,493)	
	It is(n't) true that the dots are(n't) red.	97.9	169	899 (920)	1,095 (1,089)	1,328 (1,258)	1,354 (1,427)	1,607 (1,596)	1,772 (1,765)	
Previous sentence-first experiments										
				TA (k)	FA (k + 1)	FN (k + 4)	TN (k + 5)			
Just and Carpenter, 1971 (Experiment 1) Clark and Chase, 1972 (Experiment 1) Clark and Chase, 1972 (Experiment 2) Clark and Chase, 1972 (Experiment 1) Clark and Chase, 1972 (Experiment 2)	The dots are(n't) red.	99.8	116	1,308 (1,317)	1,441 (1,433)	1,793 (1,779)	1,882 (1,895)			
	The star is(n't) above the plus.	99.8	174	1,744 (1,763)	1,959 (1,938)	2,470 (2,461)	2,624 (2,636)			
	The star is(n't) above the plus.	97.4	159	1,500 (1,539)	1,728 (1,697)	2,246 (2,174)	2,269 (2,333)			
	The star is(n't) below the plus.	99.7	170	1,875 (1,868)	2,035 (2,038)	2,520 (2,547)	2,739 (2,717)			
	The star is(n't) below the plus.	97.7	139	1,681 (1,697)	1,838 (1,835)	2,319 (2,252)	2,337 (2,391)			

				TA (k)	FA (k + 1)	FN (k + 2)	TN (k + 3)
Gough, 1966 (Experiment 2)	The boy is (n't) hitting the girl.	97.4	141	870 (861)	1,010 (1,002)	1,100 (1,143)	1,310 (1,284)
Previous picture-first experiments							
Just and Carpenter, 1971 (Experiment 2)	The dots are (n't) red.	98.0	215	1,068 (1,046)	1,255 (1,261)	1,422 (1,476)	1,728 (1,690)
Trabasso, Rollins, and Shaughnessy, 1971 (Experiment 9)	(Not) orange ^b	98.1	97	974 (978)	1,069 (1,075)	1,197 (1,172)	1,254 (1,269)
McMahon, 1963	John is (n't) slapping the girl.	97.4	155	1,060 (1,045)	1,160 (1,200)	1,390 (1,355)	1,500 (1,510)
Clark and Chase, 1972 (Experiment 2)	The star is (n't) above the plus.	99.2	272	1,783 (1,813)	2,130 (2,084)	2,354 (2,356)	2,614 (2,628)
Clark and Chase, 1972 (Experiment 3) (Attend to "both")	The star is (n't) above the plus.	99.2	341	1,911 (1,914)	2,231 (2,255)	2,652 (2,595)	2,906 (2,936)
Clark and Chase, 1972 (Experiment 3) (Attend to "top")	The star is (n't) above the plus.	99.7	390	1,596 (1,570)	1,927 (1,960)	2,337 (2,349)	2,758 (2,739)
Clark and Chase, 1972 (Experiment 3) (Attend to "bottom")	The star is (n't) below the plus.	98.1	295	1,808 (1,798)	2,115 (2,093)	2,311 (2,388)	2,726 (2,683)

Note. Abbreviations: TA = true affirmative, FA = false affirmative, FP = false predicate negative, TP = true predicate negative, FD = false denial, TD = true denial, FN = false negative, TN = true negative.

^a The number of constituent comparisons in the true affirmative case is denoted by *k*.

^b Trabasso et al. actually used the nonsense syllable *kov* to denote *not*.

comparisons. Although the slope decreased from 58 msec to 38 msec to 29 msec over the 15 days, the linearity persisted with practice. The model accounts for 97.7%, 97.4%, and 99.6% of the variance, respectively, as shown in Table 5.

In a similar experiment, Young and Chase (Note 2) collected verification latencies on 2 days, Day 3 and Day 6, separated by 2 days of practice on a related task. On both Days 3 and 6, latencies increased linearly with the number of constituent comparisons; at the same time, the slope for Day 6 was much lower than for Day 3, as shown in Table 5. We have no model of practice effects—why people get faster with extended practice. However, extended practice does seem to affect the duration of a find-and-compare operation similarly in all four information conditions. Thus, these data provide further support for a single underlying operation,

as specified by the constituent comparison model.

Chronometric Models of Sentence Verification

The constituent comparison model shares a family resemblance with a previous model proposed by Chase and Clark and Trabasso, Rollins, and Shaughnessy, hereafter the CACTRAS model (Chase & Clark, 1972; Clark & Chase, 1972; Trabasso et al., 1971). However, the resemblance ends at the comparison stage. The CACTRAS model posits a different comparison operation for each kind of constituent, rather than a single iterative operation. For example, to process a sentence representation with two constituents, CACTRAS would posit one operation that compares the inner propositions and a separate operation that compares the polarity markers. A mismatch in either comparison operation consumes

TABLE 5
SUMMARY OF THE PREDICTIONS AND FIT OF THE CONSTITUENT COMPARISON MODEL

Practice effects								
Reference	Example of a stimulus sentence	Day	Percentage of variance accounted for by model	Time per comparison (msec)	Observed (and estimated) latency			
					TA (k) ^a	FA (k + 1)	FN (k + 4)	TN (k + 5)
Singer, Chase, Young, and Clark, Note 1 ^b	The star is(n't) above.	1-5	97.7	58	635 (637)	703 (695)	841 (869)	949 (927)
		6-10	97.4	38	495 (488)	523 (526)	621 (640)	693 (678)
		11-15	99.6	29	459 (459)	488 (488)	580 (575)	598 (604)
Young and Chase, Note 2 ^b (Control condition)	The star is(n't) above the plus.							
	Subject 1	3	99.4	135	843 (831)	943 (966)	1,400 (1,371)	1,486 (1,506)
		6	97.8	46	729 (746)	814 (792)	929 (930)	971 (976)
	Subject 2	3	98.6	154	800 (783)	900 (937)	1,457 (1,399)	1,515 (1,553)
		6	99.6	65	713 (709)	771 (774)	957 (969)	1,043 (1,034)
	Subject 3	3	97.5	100	800 (793)	871 (893)	1,243 (1,193)	1,257 (1,293)
		6	96.5	32	657 (670)	714 (702)	814 (798)	814 (830)

Note. Abbreviations: TA = true affirmative, FA = false affirmative, FN = false negative, TN = true negative.
^a The number of constituent comparisons in the true affirmative case is denoted by *k*.
^b We are indebted to Bill Chase for providing the graphs from which the data points were estimated.

time, and in general, the mismatch times are different for each of the two operations.

According to the CACTRAS model, mismatches take longer than matches because the comparison itself takes longer and because changing the response index consumes time. Additionally, some of negation time may be attributed to extra time spent in reading and encoding a negative. This account fails to explain why the predicate negatives and denials in Experiment 1 have different negation times. Both types of negatives involve reading and encoding the word *not*; both involve a mismatch between an AFF and a NEG; and both require a change in the response index. If these are the processes that account for mismatch time, the negation times should be similar in the two cases. Yet, negation time is twice as long for denials as for predicate negatives. By contrast, the relative durations of the two negation times are correctly predicted by the constituent comparison model.

Although there are some difficulties with the CACTRAS explanation of mismatch time, the model can still be extended to account for the difference between predicate negatives and denials. The CACTRAS model would have to posit two different comparison operations for the two kinds of negatives. Therefore, there would be a total of three different comparison operations, one for each of the constituents: inner propositions, inner polarity markers, and outer polarity markers. The mismatch times for these three operations would be estimated as 254, 373, and 779 msec, respectively. These parameters account for 98.3% of the variance among the six conditions. By contrast, the constituent comparison model's one parameter, representing the time (200 msec) to find and compare a pair of constituents, accounts for 97.7% of the variance. The two additional parameters in the CACTRAS model account for an additional but insignificant .6% of the variance, $F(1, 115) = 1.00$.

We can also compare the fit of the present model to the CACTRAS model for the eight experimental conditions of Clark and Chase listed in Table 4. Using two parameters

for each condition, the CACTRAS model accounts for 99.8%, 97.6%, 99.8%, 98.6%, 99.5%, 99.8%, 99.7%, and 99.3% of the variance in the eight conditions. The constituent comparison model accounts for 99.8%, 97.4%, 99.7%, 97.7%, 99.2%, 99.2%, 99.7%, and 98.1%, respectively, and with only one parameter. Thus, the more parsimonious constituent comparison model is almost as precise at estimating these latencies.

The use of response latencies to infer durations of processing stages requires that certain assumptions be met. Since the CACTRAS model posited separate comparison operations for each kind of constituent, it was necessary to demonstrate that a mismatch between the polarity markers consumes the same amount of time, regardless of whether the inner propositions match, and vice versa. In fact, the data showed that the additivity assumption was quite reasonable. This additivity can be reinterpreted in terms of the constituent comparison model's single iterative find-and-compare operation. According to the current model, the two operations in the comparison stage of CACTRAS are actually multiple executions of the same operation, and so naturally, their durations are additive.

THE CONSTITUENT COMPARISON MODEL APPLIED TO RECODING

The introspective reports and the verification latencies from a number of previous studies indicate that there is another way to process negative sentences. The alternative strategy is to recode either the sentence or picture representation before comparing the two. Recoding refers to the process of changing a representation from one form into some other equivalent form. For example, a negative sentence can sometimes be recoded into an affirmative (e.g., *The door is not open* can be recoded as *The door is closed*). The result is that the recoded sentence representation is affirmative and matches the polarity of the picture representation. In this section, we will propose a recoding model based

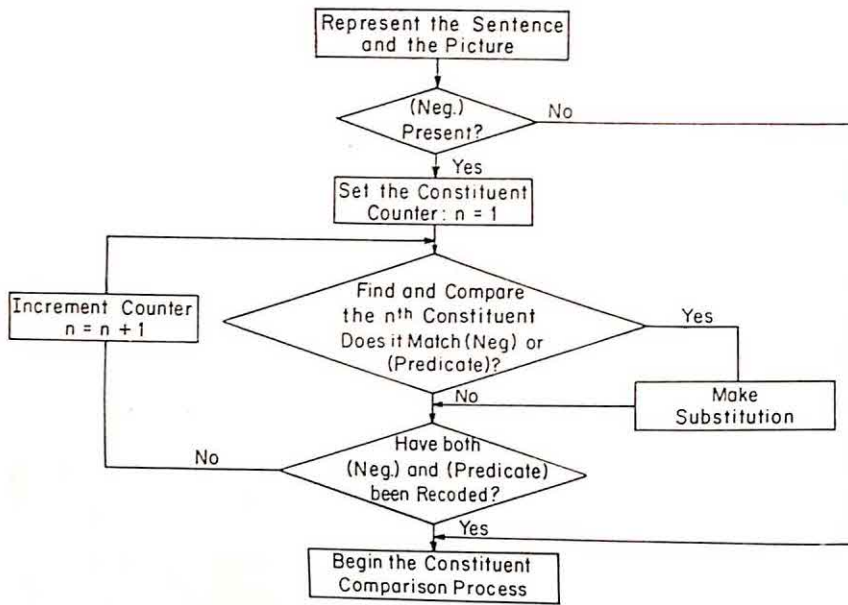


FIGURE 5. The constituent recoding model.

on the same processes as the main verification model.

Our examination of previous recoding experiments indicated that recoding a constituent consumes about the same amount of extra time as a mismatch on that constituent. Therefore, recoding time may be proportional to the number of find-and-compare operations necessary to recode the constituent. Using this as a starting point, we will attempt to account for the recoding process in terms of the same find-and-compare operation that is central to the constituent comparison model.

The recoding model (shown in Figure 5) is best explained by working through an example. In one recoding study, Young and Chase (Note 2) explicitly instructed subjects to recode negative sentences like *The star isn't above the plus* into an affirmative form with the same preposition, like *The plus is above the star*. The subjects were timed while they read the sentence, recoded it if it was negative, examined a picture like $\frac{+}{+}$ or $\frac{+}{-}$, and responded true or false. Three highly practiced subjects participated in the experiment.

The negative sentences are represented as denials (as shown in Table 6) because the ratio of negation time to falsification time was 4:1 in the nonrecoding condi-

tion. After the sentence is initially represented, it is tested for the presence of a negation marker. If the negation marker is absent, as it is for affirmatives, then processing continues as specified by the constituent comparison model. The presence of a negation marker triggers the recoding process, which consists primarily of finding the constituents to be recoded. Each successive constituent is retrieved and tested to determine whether it is to be recoded. If it is, the appropriate alternative is substituted. The actual substitution is assumed to take a negligible amount of time. If the constituent is not one of those to be recoded, then the next constituent is retrieved and examined. This recoding process terminates when all the to-be-recoded constituents have been replaced. The duration of the recoding process is therefore proportional to the locus of the last constituent to be recoded. After recoding is completed, the recoded sentence representation is compared to the picture representation as specified by the constituent comparison model. Since the recoded representation is now affirmative, the true condition will entail fewer operations than the false condition. The total latency is proportional to the sum of recoding plus comparison operations. The

TABLE 6
REPRESENTATIONS AND PREDICTIONS OF THE MODEL FOR THE YOUNG AND CHASE RECODING EXPERIMENT

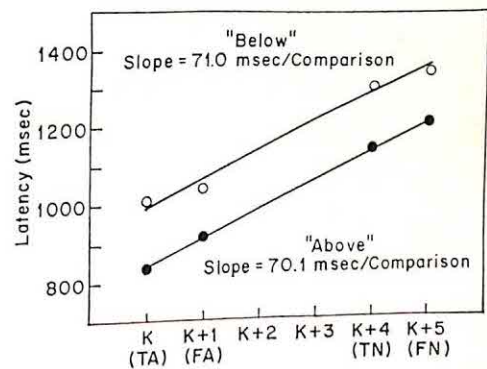
Stimulus and representation	True affirmative	False affirmative
Picture	* +	+
Sentence	The star is above the plus.	The star is above the plus.
Sentence representation	[AFF, (STAR, TOP)]	[AFF, (STAR, TOP)]
Recoding operations	(none—begin comparison process)	(none—begin comparison process)
Picture representation	[AFF, (STAR, TOP)]	[AFF, (PLUS, TOP)]
Comparison operations	+ + response = true k operations	- index = false + + response = false $k + 1$ operations
	True negative	False negative
Picture	+	*
Sentence	The star isn't above the plus.	The star isn't above the plus.
Sentence representation	{NEG, [AFF, (STAR, TOP)]}	{NEG, [AFF, (STAR, TOP)]}
Recoding operations	+ + +	+ + +
Recoded sentence	{AFF, [AFF, (PLUS, TOP)]}	{AFF, [AFF, (PLUS, TOP)]}
Picture representation	(PLUS, TOP)	(STAR, TOP)
Comparison operations	+ + + response = true $k + 4$ operations	- index = false + + response = false $k + 5$ operations

model predicts a total of $k + 4$ operations for the true negative, 3 operations for recoding, and $k + 1$ comparisons as shown in Table 6. For false negatives, it predicts a total of $k + 5$ operations, 3 for recoding, and $k + 2$ for comparison. The predictions for affirmatives remain the same, since they are not recoded: true affirmatives should take k operations; and false affirmatives, $k + 1$ operations.

The data for one subject are shown in Figure 6, along with the predictions of the model. The slope of the best fitting straight line is 70.1 msec/operation for *above*; 71.0 msec/operation for *below*. The model accounts for 99.8% and 98.4% of the variance in these two conditions. These data support the assumption that the extra time for recoding a particular constituent is equal to the time for processing a mismatch on that constituent. The data for the other two subjects are given in Table 7.

There are two differences between this recoding model and the main constituent comparison model. First of all, the recoding process is self-terminating, since there

is no need to retrieve further constituents once the appropriate alternatives have been substituted. A second difference is that the retrieval process is not reinitialized after each substitution operation. One possible reason for this latter difference is that the response index is not



NUMBER OF RECODING AND COMPARISON OPERATIONS

FIGURE 6. The fit of the constituent recoding model to data from a subject who was instructed to recode negative sentences (data from Young & Chase, Note 2). (Abbreviations: TA = true affirmative, FA = false affirmative, TN = true negative, FN = false negative.)

TABLE 7
SUMMARY OF THE PREDICTIONS AND FIT OF THE CONSTITUENT COMPARISON MODEL

COMPARISON MODEL									
Reference	Example of a stimulus sentence	Percentage of variance accounted for by model	Time per comparison (msec)	Observed (and estimated) latency					
Previous recoding experiments									
				TA (<i>k</i>) ^a	FA (<i>k</i> + 1)	TN (<i>k</i> + 4)	FN (<i>k</i> + 5)		
Young and Chase, Note 2 ^b (Transformation 2)	The star is(n't) above the plus. Subject 1	99.8	70	843 (843)	914 (913)	1,114 (1,123)	1,200 (1,193)		
	Subject 2	96.6	83	743 (732)	814 (815)	1,014 (1,063)	1,186 (1,146)		
	Subject 3	98.9	61	800 (794)	843 (855)	1,057 (1,038)	1,086 (1,099)		
Young and Chase, Note 2 ^b (Transformation 2)	The star is(n't) below the plus. Subject 1	98.4	71	1,014 (998)	1,043 (1,069)	1,300 (1,282)	1,343 (1,353)		
	Subject 2	97.1	72	843 (839)	915 (911)	1,086 (1,126)	1,229 (1,198)		
	Subject 3	98.8	62	814 (816)	886 (878)	1,043 (1,065)	1,143 (1,127)		
Slobin, 1966 (adults)	The dog is(n't) chasing the cat.	98.5	198	930 (918)	1,120 (1,116)	1,270 (1,314)	1,540 (1,512)		
Current recoding experiment									
				TA (<i>k</i>)	FA (<i>k</i> + 1)	TP (<i>k</i> + 2)	FP (<i>k</i> + 3)	FD (<i>k</i> + 4)	TD (<i>k</i> + 5)
Experiment 2, Blocks 2-7	It is(n't) true that the dots are(n't) red.	98.8	118	833 (835)	957 (953)	1,099 (1,071)	1,143 (1,189)	1,310 (1,307)	1,437 (1,425)
Note. Abbreviations: TA = true affirmative; FA = false affirmative; TP = true positive; FP = false positive; FD = false detection; TD = true detection.									

Note. Abbreviations: TA = true affirmative, FA = false affirmative, TN = true negative, FN = false negative, TP = true predicate negative, FP = false predicate negative, FD = false denial, TD = true denial.
^a The number of constituent comparisons in the true affirmative case is denoted by k .
^b We are indebted to Bill Chase for providing the graphs from which the data points were estimated.

involved in the recoding operation. The simulation suggested that items and operations involved in changing the response index were responsible for reinitializing the find-and-compare operations. Since recoding does not involve these extra items and operations, the iterative retrieval process does not have to be reinitialized.

Experiment 2: Blocks 2-7

The recoding model can be applied to the verification of the embedded negative sentences in Experiment 2, discussed earlier. In a series of pilot studies in which the sentence preceded the picture, true predicate negatives were verified faster than false predicate negatives after the subjects were practiced. This result was obtained for Blocks 2-7 of Experiment 2, indicating that subjects were spontaneously recoding predicate negatives after one

block of trials. Therefore, the data from the last six blocks of trials of Experiment 2 were analyzed separately to test the recoding model.

The recoding of predicate negatives would proceed much as we specified in our analysis of the Young and Chase study. For example, suppose the predicate negative sentence *It's true that the dots aren't red* was represented as [NEG, (RED, DOTS)] and the picture contained green dots. The color predicate *red* would be replaced by *green* and then the negation marker would be replaced by an affirmative marker, after two find-and-compare operations. After recoding, the sentence representation is affirmative, so there would only be a base number of comparisons, k , between the sentence and picture representations, for a total of $k + 2$ operations. In the false predicate negative

condition, the recoding plus comparison operations result in a total of $k + 3$ operations. In summary, when a recoding strategy is used, true predicate negatives should involve one less operation than false predicate negatives, and consequently, should be verified faster. The predicted number of operations for all six conditions are true affirmative, k ; false affirmative, $k + 1$; true predicate negative, $k + 2$; false predicate negative, $k + 3$; false denial, $k + 4$; and true denial, $k + 5$. Recoding was only initiated with the onset of the picture which provided the color predicate to be used in recoding. Therefore, the latencies should reflect the duration of both recoding and comparison operations.

Results and discussion. The mean latencies for correct responses in Blocks 2 to 7 of Experiment 2 are presented in Figure 7. The abscissa still represents an increasing number of operations—both recoding and comparison operations, although true predicate negatives are now hypothesized to involve fewer operations than the false ones. As before, the latencies increased linearly with the hypothesized number of operations. The slope of the best fitting straight line indicates that it takes about 118 msec (standard error = 20 msec) for each additional constituent comparison. Practice effects account for this slope being slightly (but not significantly) smaller than the 169 msec slope found in only the first block. The model accounts for 98.8% of variance among the six means, $F(1, 115) = 127.35$, $p < .01$. The residual 1.2% is not significant $F(4, 115) < 1$, and the root mean squared deviation was a small 23 msec. The error rates for the six conditions were low (2.8% overall) and correlated ($r = .93$) with the latencies as shown in Figure 7.

The variances in each of the six conditions also provide information that is relevant to the predictions of the model. The latencies in the six conditions are postulated to include the time for multiple iterations of a single operation. Therefore, the six variances should increase linearly with the increasing number of iterations. In fact, a linear trend accounts for 89%

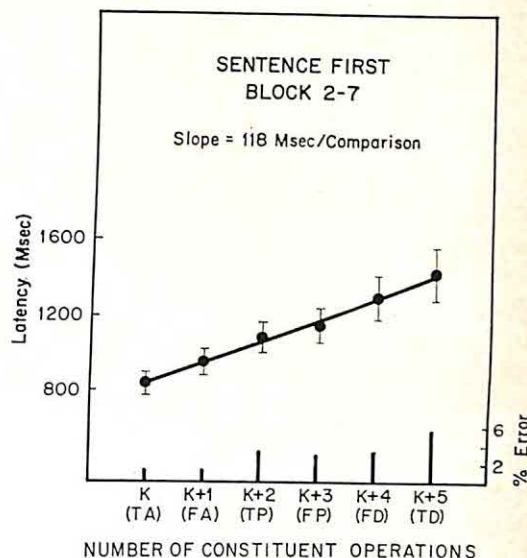


FIGURE 7. The fit of the constituent recoding model to the results of Blocks 2-7 of Experiment 2. (Abbreviations: TA = true affirmative, FA = false affirmative, TP = true predicate negative, FP = false predicate negative, FD = false denial, TD = true denial.)

of the variance among the six variances in the present experiment, 77% in Experiment 1, and 78% in Block 1 of Experiment 2, consistent with the hypothesis of a single underlying comparison operation.

The results for Experiments 1 and 2 show that predicate negatives and denials differ in two somewhat independent ways. First, predicate negatives have a smaller scope of negation, hence they take less time to process. Second, predicate negatives, unlike denials, are sometimes spontaneously recoded. The reason for this difference is unclear; perhaps it is because the recoded constituents are adjacent in the internal representation of the predicate negative but not in the denial. Of course, subjects can recode a denial when explicitly instructed, as in the Young and Chase (Note 2) study.

It should be pointed out that there are alternative recoding strategies besides the one described above. However, most of them, like recoding the picture representation into a negative form when it follows a predicate negative sentence, make the same latency predictions. Still another possibility is to recode a predicate negative

into an affirmative disjunction, like *It's true that the dots are either black or green*. This latter strategy is implausible since subjects did not report using it, and it is difficult to train someone to adopt it.

What conditions encourage people to adopt a recoding strategy? For one thing, recoding is more likely if the constituent to be recoded has a contradictory antonym, like *open* and *closed* or *odd* and *even*, where the negation of one is equivalent to the affirmation of the other. A second condition that encourages recoding is a delay between the presentation of the sentence and the second source of information (Carpenter, 1973; Trabasso, 1972). The delay may provide time for subjects to decide whether the sentence is to be recoded, to retrieve the alternative constituents, and perhaps to perform the recoding. If these conditions are satisfied, well-practiced adult subjects will often recode spontaneously. And of course, a sure way to get people to recode is to explicitly tell them how and when to do so.

Recoding in Previous Studies

We can look for evidence of recoding in a number of previous studies, particularly when the sentence was presented substantially before the picture. We can infer that recoding occurs wherever true negatives are verified faster than false negatives. For example, Slobin (1966) did a developmental study of sentence verification, in which the sentence preceded the picture and timing began at the end of the sentence presentation. For adults, true negatives were verified faster than false negatives, implicating a recoding strategy. Since the ratio of negation time to falsification time was 2.2:1, the model predicts the following number of operations: true affirmative, k ; false affirmative, $k + 1$; true negative, $k + 2$; and false negative, $k + 3$. This model accounts for 98.6% of the variance among the four means for adults. This data is shown in Table 7. The results for children, on the other hand, indicated that they did not recode these same negative sentences. The non-recoding model accounts for 97.2%, 97.7%,

94.5% and 91.1% of the variance for the 6-, 8-, 10-, and 12-year-olds, respectively. Thus, under the very same stimulus conditions, only adults seem to adopt a recoding strategy.

It sometimes happens that the latencies for true negatives and false negatives are equal, implying that recoding occurred approximately half of the time (e.g., Gough, 1965; 1966, Experiment 1). Such a mixture of recoding and nonrecoding strategies would predict k operations for true affirmatives, $k + 1$ operations for false affirmatives, and on the average, $k + 2.5$ operations for true and false negatives. In Gough's studies, recoding could only be initiated with the onset of the picture which provided the predicates to be substituted. Therefore, in both studies, the clocked interval should reflect the duration of both recoding and comparison operations. In fact, the predicted number of operations account for 99.8% and 98.9% of the variance in Gough (1965) and Gough (1966, Experiment 1), respectively. This indicates that even results produced by a mixture of strategies can be accounted for by the model.

We have argued that the find-and-compare operation in the recoding process is the same as the find-and-compare operation in the constituent comparison process. This view was supported by the ability of the basic model to account for verification latencies even when recoding strategies were used. The commonality of this constituent-by-constituent find-and-compare process suggests that it is a key operation underlying sentence comprehension.

GENERALITY OF THE CONSTITUENT COMPARISON MODEL: EXTENSION TO OTHER SEMANTIC STRUCTURES

Even though the model can very elegantly account for the processing of explicit negatives, it is also applicable to a wider range of semantic structures. To demonstrate that the model is really a more general theory of comparison processes, we will examine a number of previous studies that investigated other se-

mantic structures like counterfactuals, implicit negatives, passives, and quantifiers to see how the model fares in predicting verification latencies.

Counterfactuals. The counterfactual construction describes a situation that didn't occur, for example, . . . *if Judy had left*. While the counterfactual conditional has additional meaning components as well, at least one component of a counterfactual clause seems aptly described by some kind of negation, in this case, a negation of Judy's leaving. Therefore, this construction should be amenable to the analysis given explicit negatives. Carpenter (1973) presented sentences containing either a factual clause, (e.g., . . . *since Judy had left*) or a counterfactual clause (e.g., . . . *if Judy had left*), and then presented a probe sentence (e.g., *Judy left*) to be verified as true or false. The counterfactual clause might be represented as [NEG, (LEFT, JUDY)] and the probe sentence as [AFF, (LEFT, JUDY)]. The constituent comparison model predicts a linear increase in latencies from true factuals, to false factuals, to false counterfactuals, to true counterfactuals. In fact, the model accounts for 99.8% of the variance among the four means (as shown in Table 8). This supports the hypothesis that one component of a counterfactual is processed much like an explicitly negative clause. A similar approach may apply to several other modal constructions that have a negative component like *Mary should have*

left, Mary couldn't leave, and Mary wouldn't leave. Verification of the negative component in counterfactuals and these other modal constructions can be accounted for by the constituent comparison model without any additional assumptions.

Negative quantifiers. Negation does not always have to be explicit and use the *not* morpheme. Often, negation is implicit and lexically incorporated into an adverbial. For example, it can be demonstrated by using Klima's (1964) co-occurrence criteria that words like *few* and *seldom* are syntactically negative. Clauses with these terms have the same syntactic properties as comparable clauses with *not*. Therefore, the model might be able to explain the processing of these implicit negatives.

Just and Carpenter (1971) timed people as they verified sentences that contained negative quantifiers like *Scarcely any (hardly any, or few) of the dots are red*. Subjects were given a sentence that was negative, like one above, or a corresponding affirmative like *Many (lots of, most of) the dots are red*. The picture that followed the sentence was composed of an array of dots, containing a large subset and a small subset where one subset was black and one red. Since the ratio of negation time to falsification time was approximately 2:1, these implicit negatives must have been represented with a small scope. The data were collapsed over the three exemplars of each type of quantifier. The model predicts that latencies should in-

TABLE 8
CONSTITUENT COMPARISON MODEL APPLIED TO OTHER SEMANTIC STRUCTURES

Reference	Example of a stimulus sentence	Percentage of variance accounted for by model	Time per comparison (msec)	Observed (and estimated) latency			
				TA (k) ^a	FA ($k + 1$)	FN ($k + 2$)	TN ($k + 3$)
Counterfactuals Carpenter, 1973	If Mary would have left . . .	99.8	355	2,264 (2,281)	2,647 (2,635)	3,016 (2,990)	3,323 (3,344)
Negative quantifiers Just and Carpenter, 1971, Experiment 1	Few of the dots are red.	96.3	159	1,491 (1,516)	1,690 (1,675)	1,883 (1,834)	1,954 (1,993)
Passives Gough, 1966, Experiment 2	The car was(n't) hit by the truck.	97.8	99	1,020 (1,019)	1,130 (1,118)	1,190 (1,217)	1,330 (1,316)
Semantic memory Wason and Jones, 1963	Seven is(n't) an even number.	92.1	226	1,660 (1,741)	2,080 (1,967)	2,210 (2,193)	2,370 (2,419)
Implicit negatives Just and Clark, 1973	If John forgot to let the dog out, then the dog is out, true or false?	94.6	285	2,814 (2,890)	3,252 (3,175)	3,536 (3,461)	3,670 (3,746)

Note. Abbreviations: TA = true affirmative, FA = false affirmative, FN = false negative, TN = true negative.

^a The number of constituent comparisons in the true affirmative case is denoted by k .

crease linearly from true affirmative to false affirmative to false negative to true negative. These predictions account for 96.3% of the variance among the four means (as shown in Table 8). The individual exemplars differ in mean latency, but the relation among the four information conditions is fairly stable. Overall, the model provides a very creditable account for the processing of implicitly negative quantifiers.

Implicit negatives. The implicitly negative predicate *forget* to presents an interesting opportunity for examining the comprehension of negation. Not only is this a negative predicate, but the proposition embedded in it is also negative. For example, the sentence *John forgot to let the dog out* directly implies that John did not let the dog out (Karttunen, 1971). Thus one can study how people extract information from the implications of implicitly negative predicates, as Just and Clark did (1973). In that study (Experiment 2), subjects were presented with an affirmative sentence (*John remembered to let the dog out*) or an implicitly negative one (*John forgot to let the dog out*) and then were timed as they verified the probe sentence (e.g., *The dog is in*) as true or false of the implication of the parent sentence. The relevant information from a sentence like *John forgot to let the dog out* is that the dog is not out. This may have been represented as [NEG, (OUT, DOG)]. The information from a sentence with *remembered to* would be represented as [AFF, (OUT, DOG)]. This sentence representation would be compared to a representation of the probe, like *The dog is in*, represented [AFF (IN, DOG)]. The model predicts that verification latencies should increase linearly from true (*remembered*) to false (*forgot*) to true (*forgot*). The data conform very nicely to predictions of the model, which account for 94.6% of the variance among the four conditions as shown in Table 8. Elsewhere, we have detailed how this model can also explain the processing of other implicit negatives like *different*, *con-*

flict, *unrelated*, and *except* (Just & Carpenter, in press).

While we cannot specify a priori whether or not a word is internally represented as a negative, our model does provide a valuable litmus test. If the verification latencies for a suspect word are shorter when it is false than when it is true, then the word is probably being processed as a negative in that situation. The model suggests that *forget*, *different*, *unrelated*, *conflict*, and *except* are internally represented with two components—an affirmative core as well as a negative polarity marker. And the same kind of analysis can also exonerate suspect words. For example, *small* is not processed as a negation of *large*; it is represented as an affirmative (Carpenter & Just, 1972; Just & Carpenter, 1971). Thus, the theory provides a procedure to discriminate negative lexical items from affirmative items.

Active and passive sentences. Passive sentences have been the object of considerable psychological interest (Blumenthal, 1967; Johnson-Laird, 1968; McMahon, 1963; Mehler, 1963; Miller & McKean, 1964; Olson & Filby, 1972; Savin & Perchonock, 1965; Wright, 1969), and it seems reasonable to ask how well the model accounts for the verification of passives. Gough (1966, Experiment 2) presented affirmative or negative passives, for example, *The car isn't hit by the truck*, which may have been represented [NEG, (IS HIT, CAR)], where (IS HIT) is an abstract symbol that represents being a recipient of hitting. The picture which followed showed either a car being hit by a truck or a car hitting a truck, represented as (IS HIT, CAR) or (HITS, CAR), respectively. The model predicts a linear increase in verification latencies from true affirmative passive to false affirmative passive to false negative passive to true negative passive. In fact, the line of best fit accounts for 97.8% of the variance among the four means and has a slope of 99 msec/comparison. Thus, the model accounts very well for the processing of affirmative and negative passives in this experiment.

Norwegian sentences. Are there similar

processing strategies in languages other than English, for example, languages in which the negative morpheme occurs at slightly different places in the surface structure of sentences? To explore this issue Just and Carpenter (Note 3) examined the verification of Norwegian sentences that correspond to the ones used in Experiments 1 and 2. The experiments were performed entirely in Norwegian with native Norwegian speakers as participants. In English, the negative morpheme generally occurs after a copula verb, in both subordinate and superordinate clauses (e.g., *It is not true that the dots are red* and *It is true that the dots are not red*). In Norwegian, the negative morpheme (*ikke*) occurs after the verb (*stemmer*) in a superordinate clause but before the verb (*er*) in a subordinate clause: *Det stemmer ikke at prikkene er røde* and *Det stemmer at prikkene ikke er røde*. The constituent comparison model specifies a matching process between semantic representations that are not directly dependent on surface structure. Therefore, surface structure perturbations, such as this difference in the locus of the negatives between English and Norwegian, should not alter the basic processing of these sentences.

The results showed that the processing of these sentences in Norwegian was very similar to processing in English. Without any modifications, the constituent comparison model accounted for 97.7% of the variance in the nonrecoding situation (Block 1) and 97.9% of the variance in the recoding situation (subsequent blocks). These results indicate that the model generalizes to languages with other surface structure properties, and the processes are not peculiar to the strategies devised by English speakers dealing with English sentences.

Answering questions. To be able to answer a question, a person must compare a representation of that question to some other source of information. In this sense, the comparison process is basic to question-answering tasks as well as to verification tasks. In one kind of question-answering task, a person is given a sen-

tence and then asked a question about that sentence, for example, *The truck is hit by the car. What hits?* or *What is hit?* The representation of the sentence and question, together with the model of the comparison process, can predict whether questions will be easy or difficult to answer.

There should be more operations and hence longer latencies and more errors when the representations of the sentence and question mismatch. In fact, this is in agreement with the data that exist on answering questions about active and passive sentences (Wright, 1969) and comparative sentences (Clark, 1972). In both of these domains, it is easier to answer a question about a sentence when the sentence and question representations have similar structures, for example, when both the sentence and the question are passive. Under these conditions the comparison of corresponding constituents is facilitated. Thus, the process of comparing two representations is basic not only to sentence verification but also to answering certain types of questions.

Universal and particular quantifiers. The model has also been extended to explain the verification of sentences containing universal or particular quantifiers that refer to pictures of colored geometrical figures, for example, *All of the round figures are red* and *Some of the round figures are red*. In the proposed representations of these sentences the inner propositions were embedded in quantifiers. Just (1974) measured how long people took to verify these two kinds of sentences, examining four possible set relations between the red figures and round figures. Just (Experiment 3) applied a self-terminating version of the constituent comparison model to the data, and the model can account for 91% of the variance among the means of the eight conditions. Thus, a model of serial comparison of constituents in the sentence and picture representations can explain the processing of sentences quantified with *all* and *some*.⁴

⁴ According to the model, all of the corresponding constituents in the sentence and picture repre-

Semantic memory verification. It is important to note that the constituent comparison model is not restricted to sentence-picture comparisons but applies equally well when a sentence is compared to information retrieved from semantic memory. Just's sentence-picture verification task was analogous to Meyer's (1970) semantic memory verification task (e.g., *All/some dogs are animals*), and so performance in the two situations can be directly compared. In fact, the pattern of verification times was very similar in the two situations. Similar results in the two domains might be explained in terms of a process that is common to both tasks. The process that compares the representation of the sentence to some second source of information—be it a picture or semantic memory—is common to both tasks. Therefore, the single constituent comparison model proposed by Just (1974) can account for both sentence-picture and semantic memory verification.

Affirmative and negative sentences that refer to semantic memory are also processed as the constituent comparison model specifies. Wason and Jones (1963) asked subjects to verify sentences like *Seven isn't an even number* or *Eight is an odd number*. This task required that subjects compare the information in the sentence representation with the information retrieved from semantic memory. Just as for sentence-picture comparisons, the model predicts that the latencies for the four information conditions—true affirmative, false affirmative, false negative, and true negative—should increase linearly. This they do, with a slope of 226 msec/comparison with the model accounting for 92.1% of the

sentations are compared even if a mismatch is encountered on an early comparison. This is necessary because constituents like negative polarity markers can reverse the truth value of a statement. However, in situations where people know that there are no negations, detecting a single mismatch may be sufficient to determine that the sentence is false. In fact, people may adopt the efficient strategy of stopping the comparison process after the first mismatch in tasks that employ only simple, active, affirmative declarative sentences (cf. Just, 1974).

variance among the four means as shown in Table 8. This again illustrates how the model is not constrained to situations in which a sentence is verified with respect to an accompanying picture.

In both of these cases in which the model has been extended to verification of information from semantic memory, there has been an implicit assumption that there are processes retrieving information from semantic memory. The assumption has been that the duration of these retrieval processes is small or constant relative to the duration of the comparison process. Similarly, in sentence-picture verification, it is assumed that the duration of picture encoding is small or constant relative to the duration of the comparison process. With these assumptions, the model can account for semantic memory verification, as well as sentence-picture verification. This result indicates that the comparison processes are not specific to one situation but may be general processing strategies that are used in a variety of comprehension tasks.

The mental processes described by this model are not specific to the sentence verification paradigm but occur in a large number of situations that involve verbal comprehension (see also Carroll, 1972; Trabasso, 1972). These more general processes involve relating the information from a sentence to information from a second source, such as the listener's previous knowledge of the world. For example, in order to agree or disagree with a statement, it is necessary to compare the statement to a representation of one's own belief. In order to answer a Wh-question (e.g., *Who painted the fence*), the information provided in the question (i.e., that someone painted the fence) must be compared to previous knowledge before the interrogated constituent can be retrieved. In order to acquire new information through a verbal medium, the old information in the communication will serve as a basis to which the new information is added. The determination of which information is old can only be made if the sentence representation is

compared to previous knowledge. Elsewhere, we have demonstrated how these comparison operations also occur in following simple instructions and in various other instructional situations (Just & Carpenter, in press). Thus, the basic kinds of operations described by the model are part of a large class of comparison operations that occur very commonly when we comprehend linguistic material.

CONCLUSIONS

This analysis of the sentence verification process allows us to examine various situations and determine what features of processing remain invariant. In the three situations we examined in detail, people verified simple or embedded affirmative and negative sentences, recoded negative sentences, and verified a number of related semantic structures like implicit negatives, counterfactuals, and quantifiers. The analysis suggests that several features of the processing are common to these various tasks. First, the find-and-compare operation appears in all three situations. Moreover, this operation reappears in a number of other simple cognitive tasks, such as context recall (Sternberg, 1969a) and same-different classification tasks (Posner, 1969). Second, these basic operations seem to be performed serially since the latencies increase with the number of comparison operations that are performed. Third, the processing rate lies within certain boundaries. The estimates for the retrieval and comparison operation were primarily between 100 and 300 msec. However, with extended practice, people can retrieve and compare sentence constituents at a very fast rate, perhaps as fast as 30 msec/operation. Fourth, the order in which constituents are compared reflects the linguistic structure of the sentence and the information structure of the task. And fifth, the general nature of the control structure remained similar, although not identical, in these tasks. That is, the constituents of two representations were compared and the outcomes determined the subsequent flow of control.

The analysis also indicates what features

of the processing vary across situations. Since an information processing system adapts itself to its current task, one would expect differences in behavior from task to task. One difference was in the control structure. For example, in the three situations examined, it was found that the structure of the task determined whether or not all the constituents were compared. Not all of the constituents need to be compared in order to recode a negative representation; comparisons proceed only until the negation marker and the predicate are recoded. However, when verifying affirmative and negative sentences, all of the constituents must be compared. Another difference lies in the rule that relates the number of mismatches to the decision of *true* or *false*. In all of the current studies, an even number of mismatches resulted in a correct *true* response. However, one can envision tasks where two mismatches would not lead to a *true* response. For example, suppose the sentence *The dots aren't red* were presented with a green triangle. Here, the correct response is "indeterminate." And, the basic incompatibility between the representations might be detected very quickly, perhaps during the picture encoding stage when an attempt is made to align the two arguments. Thus, the number of constituents that are compared and the rule that relates the number and locus of mismatches to the final response of true or false may vary from task to task. The role of the processing analysis is to specify the task structure and state how it influences the control structure of the processing system.

Sentence verification has been regarded in the past as a special kind of task, unrelated to others. The current model demonstrates that, on the contrary, the kinds of operations that occur in sentence verification are not specific to language processing. The verification task can be easily incorporated into the same explanatory framework that has been used to describe other kinds of tasks. Moreover, we have argued that the comparison process is not specific to a true-false verifica-

tion paradigm but rather occurs whenever a person answers a question, follows an instruction, or incorporates statements into his belief system. The main goal has been to show that a basic comparison operation, a simple organization of the information, and a straightforward control structure can account for a wide range of experimental results. The model should help place this aspect of language processing in the more general area of cognitive psychology.

REFERENCE NOTES

1. Singer, M., Chase, W. G., Young, R. M., & Clark, H. H. *Practice effects in the comparison of sentences and pictures* (Complex Information Processing Paper No. 184). Pittsburgh, Pa.: Carnegie-Mellon University, 1971.
2. Young R., & Chase, W. G. *Additive stages in the comparison of sentences and pictures*. Paper presented at the meeting of the Midwestern Psychological Association, Detroit, April 1971.
3. Just, M. A., & Carpenter, P. A. *The comprehension of complex sentences: A Norwegian sample* (Complex Information Processing Paper No. 260). Pittsburgh, Pa.: Carnegie-Mellon University, 1974.

REFERENCES

- Atkinson, R. C., & Shiffrin, R. M. Human memory: A proposed system and its control processes. In K. W. Spence & J. T. Spence (Eds.), *The psychology of learning and motivation* (Vol. 2). New York: Academic Press, 1968.
- Blumenthal, A. Prompted recall of sentences. *Journal of Verbal Learning and Verbal Behavior*, 1967, 6, 203-206.
- Carpenter, P. A. Extracting information from counterfactual clauses. *Journal of Verbal Learning and Verbal Behavior*, 1973, 12, 512-521.
- Carpenter, P. A., & Just, M. A. Semantic control of eye movements during picture scanning in a sentence-picture verification task. *Perception & Psychophysics*, 1972, 12, 61-64.
- Carpenter, P. A., & Just, M. A. Linguistic influences on picture scanning. In J. Senders & R. Monty (Eds.), *Eye movements and psychological processes*, in press.
- Carroll, J. B. Defining language comprehension: Some speculations. In J. B. Carroll & R. O. Freedle (Eds.), *Language comprehension and the acquisition of knowledge*. Washington, D.C.: V. H. Winston, 1972.
- Chase, W. G., & Clark, H. H. Mental operations in the comparison of sentences and pictures. In L. Gregg (Ed.), *Cognition in learning and memory*. New York: Wiley, 1972.
- Clark, H. H. Difficulties people have in answering the question "Where is it?" *Journal of Verbal Learning and Verbal Behavior*, 1972, 11, 265-277.
- Clark, H. H., & Chase, W. G. On the process of comparing sentences against pictures. *Cognitive Psychology*, 1972, 3, 472-517.
- Cooper, L. A., & Shepard, R. N. Chronometric studies of the rotation of mental images. In W. G. Chase (Ed.), *Visual information processing*. New York: Academic Press, 1973.
- Crothers, E. J. Memory structure and the recall of discourse. In J. B. Carroll & R. O. Freedle (Eds.), *Language comprehension and the acquisition of knowledge*. Washington, D.C.: V. H. Winston, 1972.
- Dansereau, D. An information processing model of mental multiplication. Unpublished doctoral dissertation, Carnegie-Mellon University, 1969.
- Fillmore, C. J. The position of embedding transformations in a grammar. *Word*, 1963, 19, 208-231.
- Gough, P. B. Grammatical transformations and speed of understanding. *Journal of Verbal Learning and Verbal Behavior*, 1965, 5, 107-111.
- Gough, P. B. The verification of sentences: The effects of delay of evidence and sentence length. *Journal of Verbal Learning and Verbal Behavior*, 1966, 5, 492-496.
- Haber, R. N., & Hershenson, M. *The psychology of visual perception*. New York: Holt, Rinehart & Winston, 1973.
- Halliday, M. A. K. Notes on transitivity and theme in English: II. *Journal of Linguistics*, 1967, 3, 199-244.
- Hornby, P. A. The psychological subject and predicate. *Cognitive Psychology*, 1972, 3, 632-642.
- Hunt, E., Marin, J., & Stone, P. *Experiments in induction*. New York: Academic Press, 1966.
- Jackendoff, R. S. An interpretative theory of negation. *Foundations of Language*, 1969, 5, 218-241.
- Johnson-Laird, P. N. The choice of the passive voice in a communicative task. *British Journal of Psychology*, 1968, 59, 7-15.
- Just, M. A. Comprehending quantified sentences: The relation between sentence-picture and semantic memory verification. *Cognitive Psychology*, 1974, 6, 216-236.
- Just, M. A., & Carpenter, P. A. Comprehension of negation with quantification. *Journal of Verbal Learning and Verbal Behavior*, 1971, 10, 244-253.
- Just, M. A., & Carpenter, P. A. Verbal comprehension in instructional situations. In D. Klahr (Ed.), *Cognition and instruction*. Potomac, Md.: Erlbaum, in press.
- Just, M. A., & Clark, H. H. Drawing inferences from the presuppositions and implications of affirmative and negative sentences. *Journal of Verbal Learning and Verbal Behavior*, 1973, 12, 21-31.
- Karttunen, L. Implicative verbs. *Language*, 1971, 47, 340-358.

- Kintsch, W. Notes on the structure of semantic memory. In E. Tulving & W. Donaldson (Eds.), *Organization of memory*. New York: Academic Press, 1972.
- Klima, E. S. Negation in English. In J. A. Fodor & J. J. Katz (Eds.), *The structure of language*. Englewood Cliffs, N. J.: Prentice-Hall, 1964.
- Lakoff, G. Pronominalization, negation, and the analysis of adverbs. In R. A. Jacobs & P. S. Rosenbaum (Eds.), *Readings in English transformational grammar*. Waltham, Mass.: Ginn, 1970.
- McMahon, L. E. Grammatical analysis as part of understanding a sentence. Unpublished doctoral dissertation, Harvard University, 1963.
- Mehler, J. Some effects of grammatical transformations on the recall of English sentences. *Journal of Verbal Learning and Verbal Behavior*, 1963, 2, 346-351.
- Meyer, D. E. On the representation and retrieval of stored semantic information. *Cognitive Psychology*, 1970, 1, 242-300.
- Miller, G. A., & McKean, K. A chronometric study of some relations between sentences. *Quarterly Journal of Experimental Psychology*, 1964, 16, 297-308.
- Newell, A. A theoretical exploration of mechanisms for coding the stimulus. In A. W. Melton & E. Martin (Eds.), *Coding processes in human memory*. Washington, D.C.: V. H. Winston, 1972.
- Newell, A. Production systems: Models of control structures. In W. G. Chase (Ed.), *Visual information processing*. New York: Academic Press, 1973.
- Newell, A., & Simon, H. A. *Human problem solving*. Englewood Cliffs, N.J.: Prentice-Hall, 1972.
- Olson, D. R., & Filby, N. On the comprehension of active and passive sentences. *Cognitive Psychology*, 1972, 3, 361-381.
- Posner, M. I. Abstraction and the process of recognition. In G. Bower & J. T. Spence (Eds.), *Advances in learning and motivation* (Vol. III). New York: Academic Press, 1969.
- Rumelhart, D. W., Lindsay, P. H., & Norman, D. A. A process model for long-term memory. In E. Tulving & W. Donaldson (Eds.), *Organization of memory*. New York: Academic Press, 1972.
- Savin, H., & Perchonock, E. Grammatical structure and the immediate recall of English sentences. *Journal of Verbal Learning and Verbal Behavior*, 1965, 4, 348-353.
- Simon, H. A. What is visual imagery? An information processing interpretation. In L. Gregg (Ed.), *Cognition in learning and memory*. New York: Wiley, 1972.
- Slobin, D. I. Grammatical transformations in childhood and adulthood. *Journal of Verbal Learning and Verbal Behavior*, 1966, 5, 219-227.
- Sternberg, S. Memory scanning: Mental processes revealed by reaction-time experiments. *American Scientist*, 1969, 57, 421-457. (a)
- Sternberg, S. The discovery of processing stages: Extensions of Donders' method. In W. G. Koster (Ed.), *Attention and Performance II*. Amsterdam: North-Holland, 1969. (b)
- Trabasso, T. Mental operations in language comprehension. In J. B. Carroll & R. O. Freedle (Eds.), *Language comprehension and the acquisition of knowledge*. Washington, D.C.: V. H. Winston, 1972.
- Trabasso, T., Rollins, H., & Shaughnessy, E. Storage and verification stages in processing concepts. *Cognitive Psychology*, 1971, 2, 239-289.
- Wason, P. C. The processing of positive and negative information. *Quarterly Journal of Experimental Psychology*, 1959, 11, 92-107.
- Wason, P. C., & Jones, S. Negatives: Denotation and connotation. *British Journal of Psychology*, 1963, 54, 299-307.
- Weinreich, U. On the semantic structure of language. In J. H. Greenberg (Ed.), *Universals of language*. Cambridge, Mass.: M. I. T. Press, 1963.
- Wright, P. Transformations and the understanding of sentences. *Language and Speech*, 1969, 12, 156-166.

(Received March 15, 1974)

Birth Order and Intellectual Development

R. B. Zajonc and Gregory B. Markus
University of Michigan

A confluence model is developed that explains the effects of birth order and family size on intelligence. Intellectual development within the family context is conceived of as depending on the cumulative effects of the intellectual environment, which, for the purposes of the model, consist primarily of the siblings' and parents' intelligence. Mutual influences, through time, on the intellectual development of the siblings are described by the growth parameter α . The confluence model predicts positive as well as negative effects of birth order, a necessarily negative effect of family size, and a handicap for the last born and the only child. The model explains several features of a large birth order study carried out on nearly 400,000 19-year olds. A number of implications of the model are discussed, among them the effects of age separation between successive children. In agreement with the implications, data on the relatively low IQ of twins and triplets are cited. Extensions of the confluence model to other social processes are discussed. The confluence model is examined for its usefulness in explicating a general class of social-psychological problems: the emergence of individual differences in a social context.

A recently published review on birth order effects (Schooler, 1972) has all but put this field of research to rest. That article, entitled "Birth order effects: Not here, not now!", concludes that "The general lack of consistent findings revealed by this review leaves real doubt as to whether the chance of positive results is worth the heavy investment needed to carry out any more definitive studies" (p. 174). While the general tone of the article and of its conclusion is that of understated skepticism, the author ends the article on a cryptic, but as it turns out, rather prophetic note:

On the other hand, I suspect that other investigators, including myself, will not be able to resist the temptation of taking a cheap bet on a long shot by collecting birth order data on their subjects as they pursue studies more central to their interests. And that is how we got into this trouble in the first place (N. B. Schachter, 1959). (p. 174)

Schooler's forecast about the continued fascination with birth order found solid

confirmation in an extensive series of studies concerned with the effects of malnutrition on the intellectual performance of children born in the Netherlands at the end of the second World War (Stein, Susser, Saenger, & Marolla, 1972). As an unanticipated by-product of these studies, Belmont and Marolla (1973) discovered a strong relationship between birth order and intellectual performance. Even the most skeptical of birth order researchers will find it difficult to ignore these findings which exhibit most interesting patterns and which are based on no fewer than 386,114 data points, the entire male population of the Netherlands who attained 19 years of age in the years 1963 to 1966. The Belmont-Marolla results are reproduced in Figure 1 which reveals the five following features: First, intellectual performance as measured by Raven Progressive Matrices declines with family size. Second, within each family size ($j \geq 2$, where j is the number of children), Raven scores also decline with birth order. Third, within each family size ($j > 2$), the last born shows a greater decline in intellectual performance than any other birth rank. For a given family size ($j > 2$), the decline from the j th - 1 to the j th child is about three times greater

We wish to thank Hazel Markus, Michal Bester, and David Krantz for their suggestions and critique in the preparation of this manuscript.

Requests for reprints should be sent to R. B. Zajonc, Research Center for Group Dynamics, University of Michigan, Ann Arbor, Michigan 48106.

than the average intellectual difference between adjacent siblings. Fourth, excluding last borns, intelligence and birth order are related by a quadratic function: There is a clear indication of a reduction in the rate of intellectual decline with higher birth orders, and there is an upswing for families of eight and nine. Fifth, the only child scores at about the same level as the

firstborn of a four-child family. The magnitude of these effects is revealed in the fact that the highest (firstborn of two) and the lowest (last born of nine) scores in the Belmont-Marolla data are separated by about two-thirds of one standard deviation.¹

¹ Note that in the course of transformation, a constant of over 100 was added to the original

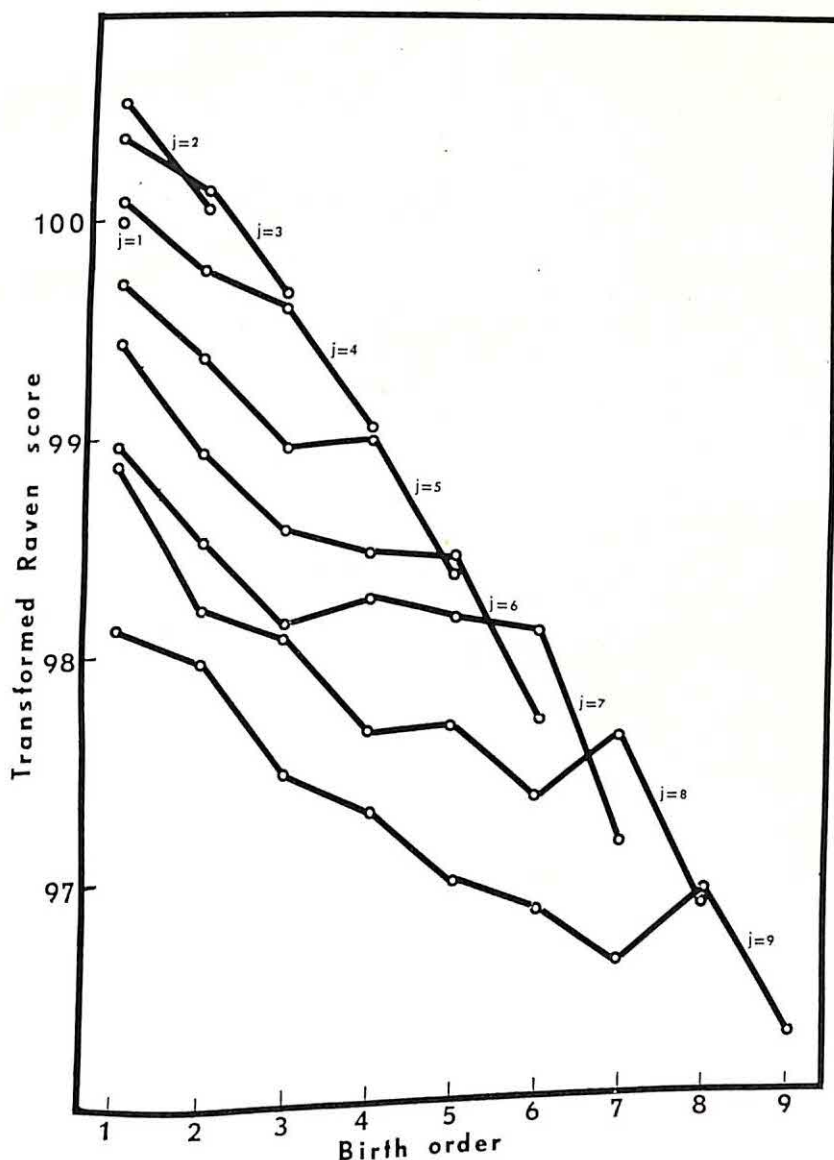


FIGURE 1. Average transformed Raven scores as a function of birth order (i) and family size (j), recalculated from Belmont and Marolla (1973). (The Raven scores were reported by Belmont and Marolla in terms of six categories from 1, high, to 6, low. For the purposes of the present analysis a linear transformation, $X_{tr} = 113.45 - 5.0047X$, was performed on these scores, inverting the scale so that increasing values now indicate increasing intelligence and setting the score of the only child at 100.)

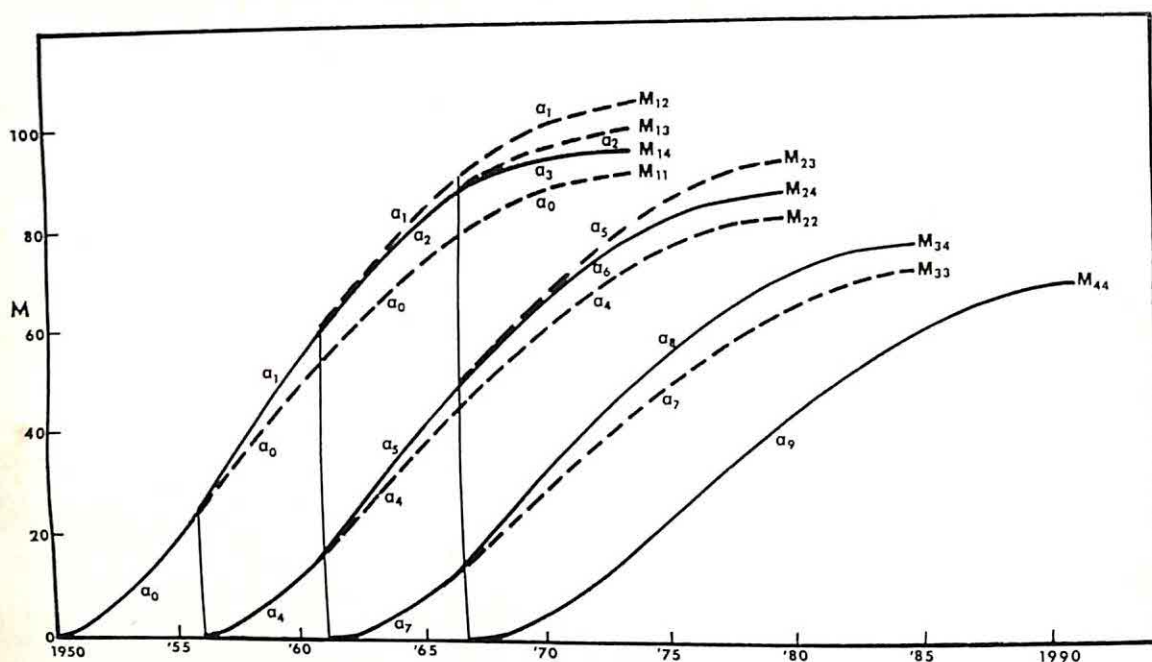


FIGURE 3. Changes in the intellectual development of siblings in a four-child family.

family is assumed to equal 100; had individual i remained an only child until adulthood, t_A , he would develop according to $f(t)$, and his intellectual level would reach $f(t_A)$. When i is t_1 years old, however, there is a change in the family configuration, and because of the changed intellectual environment, i henceforth will develop according to $g(t)$, where $\alpha_1 = 50$ in this example.

The individual's absolute intellectual level, M_{it_A} , at maturity can be expressed as the sum of the two ordinate segments m_0 and m_1 , as seen in Figure 2. The person's level at maturity cannot be estimated from $g(t_A)$ because $g(t)$ describes his growth only past the age t_1 . The intellectual increment that accrues thereafter must be added to the accumulation at t_1 . Hence, $g(t)$ must be adjusted for its intercept, and we have

$$\begin{aligned} M_{it_A} &= m_0 + m_1 = f(t) \Big|_{t_0}^{t_1} + g(t) \Big|_{t_1}^{t_A} \\ &= \alpha_0(1 - e^{-k^2 t_1^2}) \\ &\quad + [\alpha_1(1 - e^{-k^2 t_A^2}) - \alpha_1(1 - e^{-k^2 t_1^2})]. \end{aligned} \quad (2)$$

The growth parameter α represents an important aspect of this analysis, for it reflects the significant changes in the

individual's intellectual environment. But it also reflects the interdependent nature of intellectual development within the family context. Because he is part of the given environment, the individual (with his changed intellectual growth rate) affects the intellectual environment of the siblings. The various α s within the given family are, therefore, interdependent and show the mutual intellectual influences among family members. It may be noted that the later these influences come in the individual's life, the smaller will be their effect—a feature of intellectual development that is also reflected by the variable asymptote formulation. For, given $M_i = m_0 + m_1$, the older the individual, the larger proportionally is m_0 . And therefore m_1 , which denotes the proportion of growth after a change in the intellectual environment, makes less and less of a contribution.

Figure 3 shows graphically how the intellectual levels, M , of the i th child in a family of j children change as new siblings are introduced at different times. The first child is born in 1950 and grows at a rate characterized by the parameter α_0 . In 1956 a new sibling arrives, and the first-born's growth parameter changes to α_1 . Since, at that time the firstborn ceases to

be an only child, the change from α_0 to α_1 is positive. This phenomenon will be discussed fully below. In 1961 a third sibling is born, causing a decline in the intellectual environment. In 1967 a fourth sibling arrives reducing the level even further. The growth parameter in the function describing the intellectual development of the first child has thus changed from α_0 to α_1 (in 1956), to α_2 (in 1961), and finally to α_3 (in 1967). The firstborn continues to develop to maturity at the latter rate. Corresponding changes occur in the intellectual development of the other siblings in the family. Should some of the children leave home, or should the family acquire other members, for example, should grandparents join the family, the growth rates of all family members would be adjusted according to the new intellectual environment thus created.

Estimates of the α s may be obtained from information about individual adult intellectual levels and age gaps between siblings. To give an example using the aggregate Belmont-Marolla data, we know that the transformed Raven scores for M_{11} (the only child) and for M_{12} (firstborn

of two) were 100.00 and 100.57, respectively, when they reached 19 years of age. Assuming a k value of .1 and a mean gap between firstborn and second born of 2 years, we obtain the following simultaneous equations:

$$\begin{aligned} M_{11} &= \alpha_0(1 - e^{-(.1)^2(19)^2}) = 100.00 \\ M_{12} &= \alpha_0(1 - e^{-(.1)^2(2)^2}) \\ &\quad + [\alpha_1(1 - e^{-(.1)^2(19)^2}) \\ &\quad - \alpha_1(1 - e^{-(.1)^2(2)^2})] = 100.57, \quad (3) \end{aligned}$$

which yield the solution, $\alpha_0 = 102.780$ and $\alpha_1 = 103.383$.

Different sets of α s would be found by assuming different gaps between siblings. Moreover, with true individual level data on intelligence and sibling age separation, a least squares estimation procedure would be employed.

The precise nature of the relationship between the growth parameter α and the pattern of gaps can be determined only if empirical information is available about adult intelligence levels of a large population for which gaps are also known. On the theoretical level, however, the confluence model predicts different birth order effects for different gaps. Figure 4

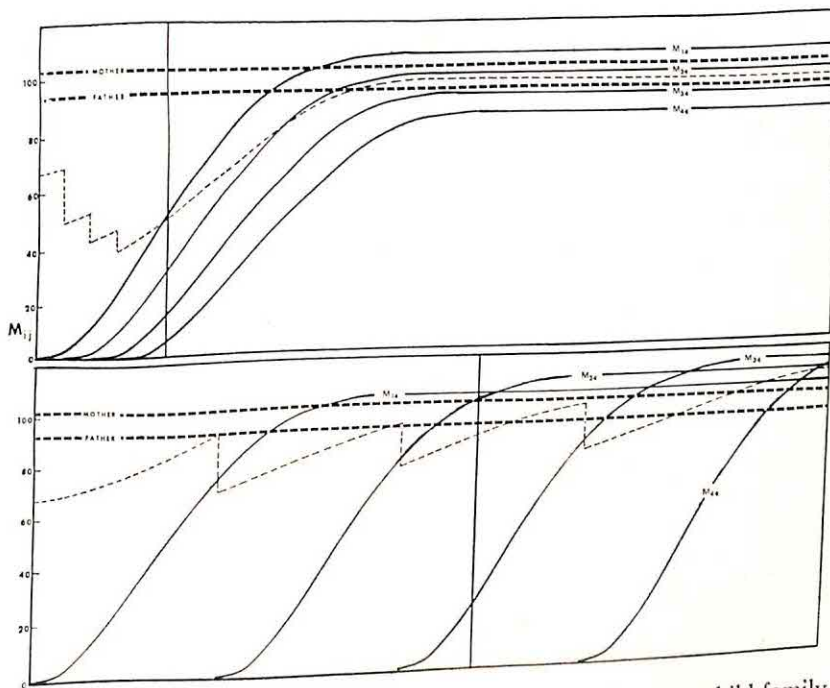


FIGURE 4. Intellectual environment (thin broken line) in a four-child family with small age separations (above) and large age separations (below).

Each graph in Figure 4 has been intersected at a point on the abscissa where the third born had reached some fixed age, the same in both graphs. In Figure 4, top where the gaps are small, intellectual environment at birth of this third child is approximately equal to 50. In the bottom portion of the figure, with the gaps extended, it is 85. It is also obvious in the top portion that the declines in the quality of the intellectual environments for each successive child are smaller and smaller. Had the number of children been extended and the small gaps preserved, eventually the newly born child would come into a *more* favorable environment than his immediate predecessor. Assuming 1-year and 2-year gaps, and, for illustrative purposes, setting the α parameters to be constant at 100, the environments at birth of nine successive children were calculated, and they are shown in Table 1. It is clear that these environments decline for the first seven children but that they begin to

TABLE 1

INTELLECTUAL ENVIRONMENT AT BIRTH											
Mother	Father	Intellectual level of sibling									Intellectual environment
		1	2	3	4	5	6	7	8	9	
1-year gap											
100	100	0									66.7
100	100	1	0								50.2
100	100	4	1	0							41.0
100	100	9	4	1	0						35.7
100	100	15	9	4	1	0					32.7
100	100	22	15	9	4	1	0				31.4
100	100	30	22	15	9	4	1	0			31.2
100	100	39	30	22	15	9	4	1	0		32.0
100	100	47	39	30	22	15	9	4	1	0	33.4
2-year gap											
100	100	0									66.7
100	100	4	0								51.0
100	100	15	4	0							43.8
100	100	30	15	4	0						41.5
100	100	47	30	15	4	0					42.3
100	100	63	47	30	15	4	0				44.9
100	100	76	63	47	30	15	4	0			48.3
100	100	86	76	63	47	30	15	4	0		52.1
100	100	92	86	76	63	47	30	15	4	0	56.6

improve thereafter, revealing the quadratic trend that characterizes the Belmont-Marolla data. This is evident for gaps equal to 1 year as well as for gaps equal to 2 years, although in the second case the upswing comes somewhat sooner. Also, as already noted, the absolute levels of intellectual environment are higher when the gaps are 2, rather than 1, years. Note that the environmental patterns shown in Table 1 were obtained with α held constant. These examples, therefore, illustrate the fact that temporal configurations of intellectual levels within the family are in themselves capable of explaining much of the Belmont-Marolla data. Thus, seen in that table is a quadratic decline with birth order, and a decline with family size can be readily inferred. A more complete account of the birth order data, however, requires the growth parameter, α , which serves three functions: It allows for scale-congruent numerical predictions; it can reflect the handicap of the last born; and when calculated from data having gap information, it can be used as an empirical index of growth changes that are brought about by changed family configurations.

This theoretical result that gaps between children affect their intellectual level leads to the prediction that IQ scores of twins and of triplets should be generally lower than the population average. This is in fact so (McCall, Appelbaum, & Hogarty, 1973; Mehrotra & Maxwell, 1949; Thurstone & Jenkins, 1929). In agreement with previous studies, Record, McKeown, and Edwards (1970), for instance, report an average verbal reasoning score of 95.7 for 2,164 twins and 91.6 for 33 triplets. A comparison group of single birth individuals earned an average score of 100.1. Admittedly, other factors may influence intellectual attainment of children of multiple births (Churchill, 1965; Lewis, 1963; Wilson, 1972). However, it is also a reasonable hypothesis that to the extent that a family with two nontwin siblings and a family with twins represent different intellectual environments, the intellectual growth of twins and nontwin siblings should reflect these differences. For families with

two children, the intellectual environment at birth of twins is $(100 + 100 + 0 + 0)/4 = 50$. For nontwin siblings, this environment is $(100 + 100 + 0)/3 = 67$ for the firstborn, and it is obvious that it must be higher than 50 for the second born, regardless of how short the gap and how close to zero the intelligence level of the firstborn when the second child is born. Moreover, each twin constitutes a large portion of the other's environment, and hence twins are less likely than other children to interact with their parents or other siblings.

It should also be clear from the above illustrations that the level of mental development decreases with family size. Regardless of the size of gaps, the overall intellectual environment declines as family size increases because in larger families the proportion of individuals with relatively low intellectual development must be necessarily high. Large gaps help younger children surpass older children, and they may cause a reversal in the relationship between birth order and intellect.

Let us return to last and only children. It was clear from Figure 1 that intellectual ability declines systematically with birth order. But the last child constitutes a departure from this pattern showing a decline that is substantially greater than that of the intermediate children. This discontinuity can be seen better in Table 2 in which linear regression slopes were computed for the Belmont-Marolla data. In nearly every case the decline from the j th

TABLE 2
DECLINE IN TRANSFORMED RAVEN SCORES WITH
BIRTH ORDER FOR LAST-BORN AND EARLIER-
BORN CHILDREN

Family size	1 to ($j-1$)	($j-1$) to j
2	—	.520
3	.270	.435
4	.257	.541
5	.253	.666
6	.250	.756
7	.146	1.056
8	.221	.811
9	.212	.625
Average	.230	.677

— 1 to the j th child is considerably greater than the average decline for other consecutive siblings. Because of their small numbers in the population, birth ranks below five or six are often collapsed in birth order studies. Hence, the sharp decline for the last born as well as the quadratic trend, such as seen in the Belmont-Marolla data, may have been inadvertently concealed in past birth order research. There are some indications that the decline of the last born is not limited to the Dutch data. Altus (1965), for example, reports Scholastic Aptitude Test data that fit the above pattern. In a two-child family the decline from the first to the second (i.e., last) born is 10 for males and 20 for females. But in a three-child family, the decline from the first to the second child is only 1 and 3 for males and females, respectively. But again, the decreases in Scholastic Aptitude Test scores are 31 for males and 13 for females when the second and last children are compared in three-child families.

There is another interesting finding in the Dutch data. Consistent with other research on birth order (Damrin, 1949; Maller, 1931; Schachter, 1963; Breland, Note 1) the only child does not fit the general pattern: Rather than occupying the dominant position, he seems to reach a level of intellectual ability comparable to that of the firstborn of a four-child family. Breland (1974) wrote recently:

Belmont and Marolla note that the mean score for only children does not follow a family size gradient. I have also observed this phenomenon . . . , but I have found no adequate explanation for it. That is, if scores tend to decline with both birth order and family size, why doesn't an only child follow this same rule and thus have the highest mean score of all? (p. 114)

Were we to extrapolate from the curve of the firstborns of families of various sizes, the expected transformed Raven score of the only child should have been 101.039. Hence, the score of 100.000 (see Figure 1) obtained by only children in the Belmont-Marolla study represents a decline of 1.039, similar to the declines of last children shown in Table 1.

The failure of only children to conform to the pattern of birth order results in the

Belmont-Marolla research and in other studies has never been understood. Generally, the only child has been regarded as an anomaly among firstborns (Altus, 1966; Belmont & Marolla, 1973; Schachter, 1963). It has been suggested, for example, that many only children come from parents who either cannot or will not have more children. Neither condition, however, necessarily implies that the child will be handicapped. Many parents who cannot have more children give their only offspring more attention and affection than would otherwise be the case. And it is not a foregone conclusion that those who want no more than one child have a necessarily depressing effect on its intellectual growth. This leaves parents who cannot have more than one child because of a specific physiological or somatic cause that had detrimental consequences for the intellectual development of the first child, parents whose first child was such an intellectual disappointment that they were unwilling to have another, and parents whose first child was an unwanted one to begin with. It is simply not known what proportions of only children come from these various types of parental environments, and there is no good evidence that only children differ in their personality from other children.

There is, however, another explanation—not necessarily contradictory—that accounts for the depressed levels of intellectual performance not only of only children but also of *all* last children. Note that we are treating the only child as a last born and ignore that he is also the firstborn. The particular circumstance that accompanies the intellectual growth of last and only children is the fact that they do not serve as intellectual resources to their siblings. If a child does not know how to solve a particular problem, he is not likely to ask the youngest; if he is missing a word, it is not the youngest who will supply it; if there is an ambiguity about the rules of a game or a pattern of play, it will not be resolved by the youngest. In short, the last born is not a "teacher" and neither is the only child. It would be quite

prising if the opportunity to perform such a "teaching" function did not have beneficial effects for intellectual development. It appears from data reported by Belmont and Marolla that there is a fairly high penalty for being last and that the removal of this penalty augments intellectual growth even under otherwise relatively unfavorable circumstances.

The last born's handicap is, of course, reflected by the α s which change with the changing family configuration. In fact, all changes in the intellectual environment are captured by the parameter α . Important new information may be obtained when we inquire about the way in which the parameter α relates to the manifest variables of birth order data. For example, the Belmont-Marolla results represent clear effects of at least three variables. It was noted in Figure 1 that the major factors of the Belmont-Marolla data were birth order (i) and family size (j). However, it was also evident that the last-born and only children exhibited a decrement larger than expected from the overall pattern. Hence a third variable, λ , the last born's handicap (including the only child as a last born), also had an influence on the data. Finally, it will be recalled that the Dutch data were characterized by a quadratic trend. Given these three variables (i , j , and λ) and the quadratic trend, the transformed aggregate Raven scores of the Belmont-Marolla study are accurately described by the following regression equation which accounts for 97% of the variance:

$$M_{ij} = 101.31 - .31i + .01i^2 - .37j + .48\lambda. \quad (4)$$

The variable λ is equal to 0 for last children and to 1 otherwise. Once obtained, the α s may be examined for the relationship to these variables, allowing us to evaluate the relative influence of these factors on intellectual growth within the family.

In order to obtain the α s, we posited a demographically reasonable gap pattern to be used for heuristic purposes, in conjunction with the Belmont-Marolla aggregate data. The pattern assumed an

average separation between siblings of 2 years except for the last born, who is separated from his closest sibling by 4 years. Since there are $1 + 2 + \dots + 9 = 45$ α s to be evaluated and, in the aggregate data, 45 Adult Raven scores, the growth parameters are obtained exactly by means of simultaneous equations.

With 2- and 4-year gaps, when the 45 α s are regressed upon the four independent variables in Equation 4, 69% of the total variance in the α s is accounted for. Of the four variables, family size (j) was significant at the .001 level ($t = 8.16$) and the last born's handicap (λ) at the .02 level ($t = 2.46$). Of greatest interest, however, is the fact that the contributions of i and i^2 , the two birth order variables, were not significant in this multiple regression analysis. Recall in this context that, according to the confluence model, the simple relationship between birth order and intellect may be either positive or negative, depending on the gap pattern among siblings (see also Figure 4). It is less important to know what birth rank the individual occupies than how many years separate him (or her) from older and younger siblings. Family size, however, is not an artifactual variable, and it necessarily depresses intellectual environment. As j increases, so does the proportion of family members with low absolute intellectual levels. Hence the intellectual environment is necessarily diluted. Within a given sibship, large age separations between successive siblings may partly mitigate the negative effects of family size. While it is for future studies to determine what might be the trade-off between family size and age separation, it is obvious that the counteraction of the ill effects of family size by age gaps is severely limited by the physical constraints of mother's age. The larger the sibship the greater need be the separation between successive siblings. But with the size of sibship held constant, the sum of age separations between children cannot exceed the mother's fecund years. And moreover, with large gaps the overall intellectual level in the family will suffer because each successive child must spend relatively long

periods of his growth as the youngest of the family and suffer the handicap of the "last" child.

CONCLUSIONS

The model developed in this article was based on data collected from individuals who had been born during the Dutch famine of 1944. Apparently, no ill effects of malnutrition were found from these subjects when comparisons of intellectual performance were made between them and other cohorts whose nutritional resources were satisfactory (Stein, Susser, Saenger, & Marolla, 1972). It is, nevertheless, possible that the substandard nutritional conditions under which the Belmont-Marolla subjects, many of whom are from the above sample, were reared may have intensified birth order and family size effects without changing the overall intellectual level of the entire population. It is also possible that because the Belmont-Marolla data cover the *entire* Dutch population of 19-year-olds, socioeconomic biases may have contributed to the results. However, Belmont and Marolla do break down their results by socioeconomic class and still find significant birth order and family size effects. Moreover, Breland (1974) reports that the Belmont-Marolla results which include males alone are "remarkably similar to [his] observations in a study of almost 800,000 National Merit Scholarship participants" (p. 114) of both sexes. Since these applicants presumably come from a more homogeneous socioeconomic background, they constitute added support for the Dutch findings.

Second, the data on which the present analysis was made are scores of the Raven Progressive Matrices test, which is reported to be fairly free from cultural and environmental influences. We would expect, therefore, that verbal tests would reveal patterns of results that would be even more pronounced than those of Belmont and Marolla. The data analyzed by Breland (Note 1) are verbal in nature and they seem to be "remarkably similar." Interestingly, Altus (1965) found birth order effects in verbal but not in the quantitative

scores of the Scholastic Aptitude Test given to University of California students.

With these cautions in mind, some answers to the puzzling questions of the relationship between birth order and family size on intelligence may be summarized. A parsimonious explanation of birth order and family size effects is obtained when we consider the nature of the intellectual environment within which children develop. This "intellectual environment" may be assessed simply by averaging the absolute intellectual levels in the given family. For particular cases, of course, we may wish to know the precise nature of the social interaction between the child and his siblings and parents. But for the purposes of gross analysis, it suffices to consider the interdependence among the children's intellectual levels and nothing more. This interdependence need not exclude parents, whose intellectual level may also be affected, however little, by the presence of children. Casual observations suggest that kindergarten teachers seem to regress in their verbal habits to the level of their charges, and parents of large families whose major interaction is with intellectually immature individuals may suffer a similar intellectual decrement. Note that had intellectual environment been represented not as an average but as a sum, family size would necessarily have *positive* effects on intelligence, a result that is contrary to the typical data in this area, including of course the Belmont-Marolla study. It is also counterintuitive to treat intellectual environment as a sum. An environment consisting of one person with an IQ of 150 is surely superior to one consisting of five persons each with an IQ of 30. It is nevertheless true that an environment of five adults each with an IQ of 150 is superior to an environment of just one adult with an IQ of 150, because a newly born child enjoys an average intellectual level of 125 in the first case and of 75 in the second.

This analysis leads to the conclusion that there are two primary determinants of intellectual growth, both imbedded in the individual's intellectual development:

One is the spacing between children and the other is family size. Large spacing is *beneficial* to the younger and detrimental to the older sibling, *small spacing* is less detrimental to the older but more harmful to the younger. Family size is damaging to all. In addition, the opportunity to serve as an intellectual resource is a factor that explains the only-child and the last-child syndrome. Neither has an opportunity to be a "teacher" of others.

The five features of the Belmont-Marolla data can be explained as follows: First, intellectual performance decreases *with birth order but only when there is close spacing between successive children*. Under these conditions each successive child has access to a less favorable intellectual environment. With larger gaps between children, this pattern can be arrested and even reversed. Second, according to the confluence model, intellectual ability decreases with family size because the larger the sibship the poorer the intellectual environment. Excluded from this generalization, of course, are cases where family size is increased by adoption of nearly adult children or by the inclusion of adult relatives (grandparents, aunts, uncles, etc.). However, when the family increases as a result of new births, the relationship between intellectual environment and family size will be *necessarily* negative. In contrast to their capacity of modifying birth order effects, gaps, however large, cannot overcome the negative effects of sibship size. Third, the last child shows a larger drop in intelligence scores because of a lack of opportunity to "teach." Fourth, the only child also suffers from the "last-child" handicap. Finally, the upswing in the Raven scores which were found for the lower ranks in larger families is a feature of combining exponential functions, and it occurs because lower ranks have an environment that includes fewer individuals with an intelligence level lower than their own, while this is not the case for higher birth ranks.

If the Belmont-Marolla data represent a general pattern—and Breland's (Note 1) data suggest that they may indeed do so—

then it is clear from the above analysis that if one had a choice, one would want to be the firstborn of a family of two, with the younger sibling following closely in time. *The sooner one acquires a younger sibling, the sooner one acquires the role of "teacher."* If there already happens to be a firstborn in the family, then one would want to postpone another birth for a period of time, allowing the older sibling to gain a substantial level of intellectual maturity. *If we ignore other factors that influence intellectual growth, then under the constraints of the confluence model, the parents' dilemma can be formulated. How should they select the spacing of children which is optimal (or fair) according to some criterion?* If the gap is small, the firstborn benefits because he loses the "last-child" handicap early. But, the second born is brought into an inferior environment because of his older sibling's intellectual immaturity. The second can only benefit by a longer gap. But then the firstborn is at a disadvantage because he must spend a long period of his growth as the last born. Ignoring other factors, it is possible, in principle, to select a separation between children that minimizes (or maximizes) intellectual differences. It is equally possible to select a gap that maximizes the *average* intelligence level. The dilemma may be exacerbated by the fact that a high *average* intelligence may be obtained only at the expense of large differences between siblings. These speculative implications, however, must be substantiated by empirical information about the true effects of gaps. While the confluence model developed here leads to the derivation that, other things equal, gaps are beneficial, other things are seldom equal, and in reality the greater the gap the less the time siblings spend together—not only because the older is likely to leave home sooner but also because he engages in activities in which the younger cannot share. Hence, in practice gaps have optimal lengths. Studies that relate intelligence to gaps (Cicirelli, 1967; Koch, 1954; Thurstone & Jenkins, 1929) are based on samples that are small and do not

control for family size and birth order. Hence, they do not allow clear interpretation. They do, however, show a trend that agrees with the suspicion that the relationship between gap length and intelligence is not monotone but that there is an optimal separation between children's ages.

The confluence model shows that environment, even in the restricted sense used here, may have powerful effects on the intellectual development of the individual. It goes without saying that other factors, such as genetic background and other environmental processes not considered here (such as child rearing practices, unique experiences, etc.), all contribute to that development as well. But, although *it is beyond the present means to determine just how much variation in intellectual growth is contributed by these factors and by those considered here, it is, nevertheless, worth noting that much of the birth order data can be parsimoniously explained by considering no more than absolute intelligence levels of siblings during the course of their development.*

The confluence model need not be restricted to the growth of intelligence. Other processes that develop over time may very well have a similar form. The confluence model formulates an analysis of a social process in a way that resembles a river. It is a somewhat unusual river because even after having joined the general flow, its tributaries are distinguishable and do not lose their identities. As the separate currents, one after another, join the mainstream, they affect the general flow of the river and are themselves affected by uniting with it. The main feature of the formulation is that the individual is considered to be a part of his own environment. And this environment is conceived of not as a static and stable background condition, but as one that changes over time, and one that is dynamically interdependent with its components. The individual is continually influenced by his own environment, and being thus influenced and changed, himself brings about changes in his environment by virtue of his very own change. These mutually depen-

dent changes in the individual and in the environment occur instantaneously.² This interdependence is reflected in the simultaneous equations employed in the analysis. Quantification of this process involves the assessment of the individual at each point of time along a specified dimension, for example, intellectual level, and of the ever-changing environment that emerges. What we mean by "environment" was the surrounding intelligence whose level varies as the participants enter, leave, and change. As we have seen, these temporal changes follow a dilution pattern.

To carry out this sort of analysis one requires a reliable measure of the individual quality under examination and a knowledge of changes in this quality over time. Intelligence satisfies these requirements, and hence a confluence analysis can be carried out. If growth rates or changes over time are *not* known, an analysis is still feasible if a suitable assumption is made.

What would the application of the confluence model look like in traits other than intelligence? Clearly, the relationship between birth order and personality development constitutes a more complex problem than the development of intellectual ability. One can, however, make some speculations. With regard to traits such as

² It is possible to let α vary continuously instead of having it remain constant between successive births, as we did earlier. Let α be equal to some average of the intellectual levels of all of the children and adults in the given family at a given time t . Then the rates of intellectual growth, according to the function $f(t) = \alpha(1 - e^{-k^2 t^2})$, can be expressed by the differential equation $df/dt = 2k^2 t(\alpha - f)$. For a given intellectual level, M_i , we have $dM_i/dt = 2k^2 t[\alpha(t) - M_i(t)]$. Note that the last born's handicap would have to be represented by a separate parameter. The obvious advantages of a variable α formulation are offset by the fact that because α s now vary continuously, they can no longer be evaluated in relation to birth order, family size, and other variables that affect intellectual growth within the family if repeated measures of all individuals in the sample are not available. Since there are no such birth order data at present and since longitudinal information was not available for the Belmont-Marolla results, a continuously changing α was approximated by a step function where α changed whenever family composition changed.

dependency, for example, it is a fair hypothesis that if the first child is in fact likely to be quite dependent on its mother (Sears, 1950), the access to the mother by children born later may thereby become restricted. Later-born children may, therefore, become less dependent than the firstborn because when young they learn to seek support from a greater variety of social sources, including other siblings. The development of dependency in the family context represents what seems to be a "preemptive" pattern: If earlier children are dependent, they force younger children into a situation that requires independence because the major social targets of dependence have been preempted. Other traits may show other patterns. Intelligence produces a "dilution" pattern. Affiliation, on the other hand, may be "cumulative," with each successive child having a greater variety of affiliative targets and developing, therefore, greater affiliative skills. Gaps may influence the development of affiliative behavior, since dissimilarity of ages may tend to reduce social contact. Status needs and dominance behavior are more likely candidates for a systematic relationship to birth order. If intellectual level tends to promote assertive behavior, and to the extent that earlier-born children are treated as "teachers" by younger siblings (and often given this responsibility quite explicitly by their parents), they are more likely to develop leadership skills.

Not only permanent personality traits but also temporary individual differences that characterize the individuals' relations to a group might also be analyzed as they emerge in a social context. Take group cohesiveness, which has been conceptualized traditionally as some function of the attraction of the individual members to the group. A group is formed. It has some goal. The original founders are naturally very committed to the group goal and are very attracted to each other. New members are recruited. At the time of joining, their attraction for the group is lower than that of the founding fathers. Therefore, by joining, the new member dilutes the

"cohesiveness" environment of the group. The average attraction in this group is now lower than previously, and if there is a second recruit, he shall join a "different" group than the first. And so on. We have a similar situation here as with intelligence and the family. To the extent that attraction to the group depends on all the attractions in that group in the same way that intelligence depends on intellectual environment, and if we know the rate at which attraction grows and what is the pattern of entries, we can make predictions about differences in attraction of individual members to the group and about overall group cohesiveness. The study of changing norms and of conformity in groups with changing memberships may also benefit from the use of the confluence model. In fact, Jacobs and Campbell (1961) report experimental data on the transmission of a norm over "generations" of subjects that show a dilution pattern. However, for none of these illustrations, except perhaps for the last, is it possible at present to perform an analysis such as was done for intelligence, birth order, and family size. Large sample data, together with solid information about rates of change, are required.

REFERENCE NOTE

1. Breland, H. M. Birth order, family configuration, and verbal achievement (Research Bulletin No. 72-47). Princeton, N.J.: Educational Testing Service, 1972.

REFERENCES

- Altus, W. D. Birth order and scholastic aptitude. *Journal of Consulting Psychology*, 1965, 29, 202-205.
- Altus, W. D. Birth order and its sequelae. *Science*, 1966, 151, 44-49.
- Becker, G. Visual acuity, birth order, achievement versus affiliation and other Edwards Personal Preference Schedule scores. *Journal of Psychosomatic Research*, 1965, 9, 277-283.
- Belmont, L., & Marolla, F. A. Birth order, family size, and intelligence. *Science*, 1973, 182, 1096-1101.
- Bradley, R. W. Birth order and school-related behavior: A heuristic review. *Psychological Bulletin*, 1968, 70, 45-51.
- Breland, H. M. Birth order, family size, and intelligence. *Science*, 1974, 184, 114.

- Churchill, J. A. The relationship between intelligence and birth weight in twins. *Neurology*, 1965, 15, 341-347.
- Cicirelli, V. G. Sibling constellation, creativity, IQ, and academic achievement. *Child Development*, 1967, 38, 481-490.
- Clarke, E. L. *American men of letters, their nature and nurture*. New York: Columbia University Press, 1916.
- Damrin, D. E. Family size and sibling age, sex, and position as related to certain aspects of adjustment. *Journal of Social Psychology*, 1949, 29, 93-102.
- Eisenman, R. Birth order and artistic creativity. *Journal of Individual Psychology*, 1964, 20, 183-185.
- Farina, A., Barry, H., & Garmezy, N. Birth order of recovered and nonrecovered schizophrenics. *Archives of General Psychiatry*, 1963, 9, 224-228.
- Gelfand, S. The relationship of birth order to pain tolerance. *Journal of Clinical Psychology*, 1963, 19, 406-407.
- Glass, D. C., Horwitz, N., Firestone, I., & Grinker, J. Birth order and reactions to frustration. *Journal of Abnormal and Social Psychology*, 1963, 66, 192-194.
- Green, C. E. The effect of birth order and family size on extra-sensory perception. *Journal of the Society for Psychical Research*, 1965, 43, 181-191.
- Innes, J. M., & Sambrooks, J. E. Paired-associate learning as influenced by birth order and the presence of others. *Psychonomic Science*, 1969, 16, 109-110.
- Jacobs, R. C., & Campbell, D. T. The perpetuation of an arbitrary tradition through several generations of a laboratory microculture. *Journal of Abnormal and Social Psychology*, 1961, 62, 649-658.
- Koch, H. L. The relation of "primary mental abilities" in five- and six-year-olds to sex of child and characteristics of his siblings. *Child Development*, 1954, 25, 209-223.
- Lester, D. Sibling position and suicidal behavior. *Journal of Individual Psychology*, 1966, 22, 204-207.
- Levinger, G., & Sonnheim, M. Complementarity in marital adjustment: Reconsidering Toman's family constellation hypothesis. *Journal of Individual Psychology*, 1965, 21, 137-145.
- Lewis, M. M. *Language, thought and personality in infancy and childhood*. London: Harrap Press, 1963.
- Maller, J. B. Size of family and personality of offspring. *Journal of Social Psychology*, 1931, 2, 3-27.
- McCall, R. B., Appelbaum, M. I., & Hogarty, P. S. Developmental changes in mental performance. *Monographs of the Society for Research in Child Development*, 1973, 38, (3, Serial No. 150).
- Mehrotra, S. N., & Maxwell, J. The intelligence of twins. A comparative study of eleven-year-old twins. *Population Studies*, 1949, 3, 295-302.
- Record, R. G., McKeown, T., & Edwards, J. H. An investigation of the difference in measured intelligence between twins and single births. *Annals of the Human Genetic Society*, 1970, 34, 11-20.
- Sampson, E. E. Birth order, need achievement, and conformity. *Journal of Abnormal and Social Psychology*, 1962, 64, 155-159.
- Schachter, S. *The psychology of affiliation*. Stanford, Calif.: Stanford University Press, 1959.
- Schachter, S. Birth order, eminence and higher education. *American Sociological Review*, 1963, 28, 757-768.
- Schachter, S. Birth order and sociometric choice. *Journal of Abnormal and Social Psychology*, 1964, 68, 453-456.
- Schooler, C. Birth order and hospitalization for schizophrenia. *Journal of Abnormal and Social Psychology*, 1964, 69, 574-579.
- Schooler, C. Birth order effects: Not here, not now! *Psychological Bulletin*, 1972, 78, 161-175.
- Sears, R. R. Ordinal position in the family as a psychological variable. *American Sociological Review*, 1950, 15, 397-401.
- Sletto, R. F. Sibling position and juvenile delinquency. *American Journal of Sociology*, 1934, 39, 657-669.
- Smart, R. G. Alcoholism, birth order, and family size. *Journal of Abnormal and Social Psychology*, 1963, 66, 17-23.
- Stein, Z., Susser, M., Saenger, G., & Marolla, F. Nutrition and mental performance. *Science*, 1972, 178, 708-713.
- Thurstone, L. L., & Jenkins, R. L. Birth order and intelligence. *Journal of Educational Psychology*, 1929, 20, 641-651.
- Ward, C. D. A further examination of birth order as a selective factor among volunteer subjects. *Journal of Abnormal and Social Psychology*, 1964, 69, 311-313.
- Willis, C. B. The effects of primogeniture on intellectual capacity. *Journal of Abnormal and Social Psychology*, 1924, 18, 375-377.
- Wilson, R. S. Twins: Early mental development. *Science*, 1972, 175, 914-917.

(Received April 22, 1974)

Teaching Psychology in Secondary Schools

Coauthors

RICHARD A. KASSCHAU
and
MICHAEL WERTHEIMER

Presents a historical review of the growth and development of high school psychology courses in the U.S. In so doing, the foundation is laid for an informed discussion of the philosophical concerns of the present-day psychology teacher. The pros and cons of two major perspectives, "psychology for personal adjustment" and "psychology as a discipline," are presented and a synthesis of these approaches is outlined and evaluated. This book also provides valuable information on available curriculum materials, references useful in obtaining these materials, and teacher training and certification curricula and criteria. 1974. 64 pp. \$2.50.

Please note that payment must accompany all orders amounting to \$15 or less except for those orders submitted on institutional purchase order forms.

American Psychological Association
Publication Sales Dept.
1200 17th Street, N.W.
Washington, D.C. 20036



Publication Manual

of the American
Psychological Association

Second edition, 1974. Current guide to the preparation of manuscripts for APA and related journals. This edition adds many examples, including an entire sample paper. New sections on student papers, metrication, and mathematical copy. Enlarged sections on content and organization of an article, typing the manuscript, and all aspects of APA style, including reference lists and tables. 136 pp. \$4.

Please note that payment must accompany all orders amounting to \$15 or less except for those orders submitted on institutional purchase order forms.

**American Psychological
Association**

Publication Sales Dept.
1200 17th Street, N.W.
Washington, D.C. 20036



Thesaurus of Psychological Index Terms

Contains over 4,000 terms representing an authoritative, unambiguous, and hierarchically structured vocabulary used by psychology and related disciplines. This vocabulary will facilitate precise content representation of concepts described in psychological literature and allow for more flexibility in automated search and retrieval operations. These two factors promote more effective communication of psychological concepts in original works and research, ultimately aiding the content of publications to be focused in accordance with the interest patterns of psychologists.

This **Thesaurus**, developed by the Psychological Abstracts Information Services, is divided into three sections:

The **relationship section** shows the relationships among concepts that are associated with each other. Term relationships noted are: synonymous, broader, narrower, or related.

The **alphabetical section** presents each preferred term in alphabetical order, enabling rapid identification of terms used for indexing.

The **hierarchical section** contains 17 broad content areas in which terms are presented in descending order according to breadth of concept represented. The 17 categories correspond to major classifications used in *Psychological Abstracts*.

The relationship, alphabetical, and hierarchical sections will assist the user in selecting appropriate terms for different uses.

Administrators

- determine areas of research for an organization
- assure representative expertise in personnel selection

Researchers and Authors

- find synonyms for a term representing a concept
- select terms that represent a concept
- define subject matter topics for a paper or book

Educators

- structure a curriculum
- use as a tutorial aid

Consultants

- search the literature
- aid in retrieval of information

Librarians

- provide access to psychological literature
- aid in cataloging psychology material
- aid in ordering new psychological literature

Students

- locate concepts relevant to psychology
- focus areas of study

The **Thesaurus** contains 350 pages with a hard-cover and costs \$12. Please note that all orders amounting to \$15 or less must be accompanied by full payment except for those orders submitted on institutional purchase order forms.

American Psychological Association
Publication Sales Department
1200 17th Street, N.W.
Washington, D.C. 20036



PSYCHOLOGICAL REVIEW

GEORGE MANDLER, EDITOR

JEAN M. MANDLER, ASSOCIATE EDITOR

UNIVERSITY OF CALIFORNIA, SAN DIEGO

Vol. 82, No. 2

March 1975

Articles

- On the Temporal Control of Behavior Peter Killeen 89
- Arousal, Activation, and Effort in the Control of Attention..... Karl H. Pribram and Diane McGuinness 116

Theoretical Notes

- Generation-Recognition Theory and the Encoding Specificity Principle..... Edwin Martin 150
- Features and Semantic Memory: Set-Theoretic or Network Model?..... James D. Hollan 154
- Set-Theoretic and Network Models Reconsidered: A Comment on Hollan's "Features and Semantic Memory"
Lance J. Rips, Edward E. Smith, and Edward J. Shoben 156
- Metabolic Hormones and Regulation of Feeding: A Reply to Woods, Decke, and Vasselli
Jaak Panksepp 158
- Insulin, Growth Hormone, Body Weight, and Feeding: A Reply to Panksepp
Stephen C. Woods, Elisabeth Kaestner, and Joseph R. Vasselli 165

Editorial Consultants

The editor gratefully acknowledges the efforts of the following consultants who helped review manuscripts for this issue. They are part of a large group of psychologists who have kindly agreed to assist in the editing of manuscripts for *Psychological Review*.

NORMAN ADLER

JOHN R. ANDERSON

NORMAN H. ANDERSON

A. M. COLLINS

IRVING L. JANIS

H. M. JENKINS

DAVID H. KRANTZ

HOWARD LEVANTHAL

LEAH L. LIGHT

JOHN MORTON

LLOYD NAKATANI

ROBERT G. PACHELLA

DAVID E. RUMELHART

THEODORE SARBIN

JOAN G. SNODGRASS

JOHN E. R. STADDON

PHILIP TEITELBAUM

RICHARD F. THOMPSON

ROC E. WALLEY

RICHARD E. WHALEN

FREDERICK WIGHTMAN

GEORGE WOLFORD

The *Psychological Review* publishes articles that make theoretical contributions to any area of scientific psychology. Preference is given to papers that advance theory rather than review it, and to statements that are specifically theoretical rather than programmatic. Surveys of the literature, problems of method and design, and papers that primarily report empirical findings are ordinarily not appropriate. Letters not exceeding 250 words on topics of theoretical interest should be submitted in duplicate.

Preparation of Manuscripts: Instructions for the preparation of manuscripts appear in the *Publication Manual of the American Psychological Association* (2nd ed.). Submit three copies; one copy must be the original typed copy. All parts of the manuscript must be double spaced. In order to facilitate blind reviewing, the author's name should not appear on the manuscript or abstract, but each copy of the manuscript should include a separate title page showing the title and the author's name and affiliation. Footnotes that identify the author should be typed on separate pages. Figures should be submitted as glossy prints ready for reproduction; duplicate figures may be photocopies. Metric equivalents must be given for all nonmetric measurements. Authors should keep a copy of the manuscript to guard against loss.

Abstracts: Manuscripts must be accompanied by an abstract of 100 to 175 words typed on a separate sheet of paper. Detailed instructions for preparation of abstracts appear in the *Publication Manual* or may be obtained from the Editor.

Manuscripts: Submit manuscripts in triplicate to the Editor, GEORGE MANDLER, Department of Psychology, University of California, San Diego, La Jolla, California 92037 according to instructions provided above.

Copyright and Permissions: All rights reserved. Written permission must be obtained from the American Psychological Association for copying or reprinting text of more than 500 words, tables, or figures. Permissions are normally granted contingent upon like permission of the author, inclusion of the APA copyright notice on the first page of reproduced material, and payment of a fee of \$10 per page of APA material. Permission fees are waived for authors who wish to reproduce their own articles. Address requests for permission to the Permissions Office, American Psychological Association, 1200 Seventeenth Street, N.W., Washington, D. C. 20036.

Subscriptions: Subscriptions are available on a calendar year basis only (January through December). Nonmember rates for 1975: \$15 domestic, \$16 foreign, \$3 single issue. APA member rate: \$6. Write to Subscription Section, APA.

Reprints: Authors may order reprints of their articles from the printer when they receive galley proofs.

Back Issues and Back Volumes: For information regarding back issues or back volumes write to Order Department, APA.

Microform Editions: For information regarding microform editions write to any of the following: Johnson Associates, Inc., P.O. Box 1017, Greenwich, Connecticut 06830; University Microfilms, Ann Arbor, Michigan 48106; or Princeton Microfilms, Princeton, New Jersey 08540.

Change of Address: Send change of address notice and a recent mailing label to the attention of the Subscription Section, APA, 30 days prior to the actual change of address. Undelivered copies resulting from address changes will not be replaced; second-class mail is not forwarded unless subscribers notify the local post office in writing that they will guarantee second-class forwarding postage.

Published bimonthly (beginning in January) in one volume per year, six issues, by the American Psychological Association, Inc., 1200 Seventeenth Street, N. W., Washington, D. C. 20036. Printed in the U.S.A. Second-class postage paid at Washington, D. C., and additional mailing offices.

HAROLD P. VAN COTT
Managing Editor

ANITA DEVIVO
Executive Editor and
General Manager

DEBORAH S. FELDER
Supervising Technical
Editor

MARY GRIMBALL
Technical Editor

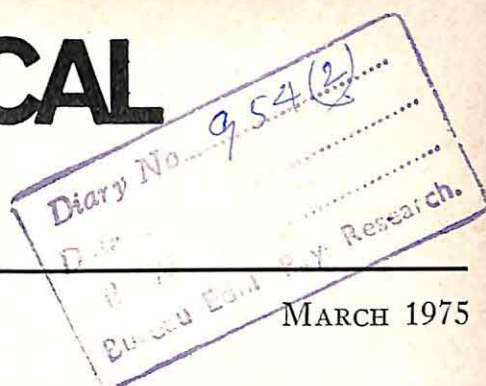
LYNN D. SPIER
Manager, Publications
Business Services

ANNE REDMAN
Subscription Manager

PSYCHOLOGICAL REVIEW

Copyright © 1975 by the American Psychological Association, Inc.

VOL. 82, No. 2



MARCH 1975

On the Temporal Control of Behavior

Peter Killeen
Arizona State University

A series of experiments ranging over a number of animal species, several classes of behavior, and a variety of experimental conditions provided the foundation for a mathematical model of the temporal control of behavior. The initial research explored the general activity of pigeons during the interval between food deliveries; most of the variance in that data was accounted for by the algebraic sum of two exponential functions. Parameters of the model were interpreted as indices of arousal level and inhibitory processes which detracted from that level. The parameters changed in ways appropriate to those interpretations when behavior was modified by the action of various drugs, reinforcement contingencies, and competing behaviors. The model was mapped onto a series-latency mechanism and extended to account for various schedule-induced, operant, and respondent behaviors.

Perhaps the most basic control procedure in the study of conditioning is the periodic presentation of the reinforcer—or unconditioned stimulus—independent of the organism's behavior and in the absence of any associated stimulus change. Such procedures have given rise to a number of theoretically important discoveries. Pavlov (1960) found conditioning to occur readily under such circumstances and called the process temporal conditioning because of the inferred temporal conditioned stimulus. Skinner (1948) found that walking and other gross movements increased in frequency under such circumstances. He called the process superstitious conditioning, citing the fortuitous contingency between various responses and the periodic reinforcer. Campbell and Sheffield (1953) employed the paradigm to study the increase in spontaneous activity that was said

to occur with deprivation. They found that the increased activity of rats actually reflected a hyper-reactivity to normal stimuli; the effect of hunger "is to augment the normal activity responses to novel stimuli and to greatly augment the activity to stimuli associated with the performance of the consummatory response" (Sheffield & Campbell, 1954, p. 99). Amsel and Work (1961), Slivka and Bitterman (1966), and Zamble (1967) have also shown that the presentation of stimuli that have been paired with a reinforcer increases general activity. Such studies have contributed impetus to the recent movement away from drive-reduction theories and toward incentive-motivation accounts of behavior (cf. Bindra, 1970; Bolles, 1967; Cofer & Appley, 1967). Recent studies of autoshaping (e.g., Brown & Jenkins, 1968) are the logical extensions of this line of enquiry.

Not all activity during periodic feeding schedules is associated with stimuli occurring prior to reinforcement; hyperreactivity to stimuli occurring after reinforcement has

The author is indebted to Arizona State University for support through several summers.

Requests for reprints should be sent to Peter Killeen, Psychology Department, Arizona State University, Tempe, Arizona 85281.

also been observed. "Schedule-induced" polydipsia, wheel running, gnawing, aggression, and pica (Falk, 1972; Segal, 1972) exemplify this type of "adjunctive" activity. Staddon and Simmelhag (1971) have also reported unreinforced sequences of behavior recurring throughout the interval between feedings.

Activity has been variously taken as an *integral part* of the normal conditioning process (Sheffield, 1966), as an *adventitious by-product* of conditioning by contiguity (Skinner, 1948), as a sign of *frustration* (cf. Scull, 1973), and as an *adjunctive* behavior induced by a schedule of reinforcement (Levitsky & Collier, 1968). Most of these evaluations have been made within the context of one or another theoretical paradigm; few have incorporated extensive parametric measurements of activity; none have adduced an adequate quantitative description of the course of activity over time. The following experiments demonstrate that an increase in activity level is as fundamental a consequence of reward as is the acquisition of learned responses; that activation is not reducible to the law of effect; and that activity increases to a maximum soon after a signal of impending reward and then decreases as a function of interference by competing responses. These and other implications are drawn from an empirically derived model of activation. An extension of the model adequately describes the temporal course of typical operant and respondent behaviors and of schedule-induced behaviors such as polydipsia.

ACTIVITY AS A FUNCTION OF INTERFOOD INTERVAL

Experiment 1

Method. Four White Carneaux pigeons, all with previous histories of experimentation, served as subjects. The wooden experimental chamber was 46 cm wide, 31 cm deep, and 34 cm high. A feeder that dispensed mixed grain was centrally located on the widest wall, and a houselight was located directly above it near the ceiling. This chamber was placed in a sound-resistant enclosure, which itself was located in a room separate from the control and recording equipment. The floor of the chamber consisted of six movable panels, three hinged at the rear of the chamber and three hinged at the front. Microswitches under these panels

indicated when the animal moved from one part of the chamber to another or shifted its weight while staying in one part of the chamber. Each activation of a microswitch counted as a single "activity response." Data from the six switches were pooled with no account taken of where in the chamber a response occurred.

The pigeons were reduced to 80% of their free-feeding weight and exposed to various feeding schedules. On these schedules the food hopper was activated for 3 sec at various interfood intervals (25, 60, 120, 200, and 400 sec). Half of the animals experienced the intervals in an ascending series, and half in a descending series. To preclude adventitious reinforcement of pacing, food was delivered no sooner than 5 sec after the last recorded activity in the chamber (5-sec Differential Reinforcement of Other Behavior, or DRO). Ten daily sessions were conducted at each interfood interval. Each session lasted for 60 feedings, except at the longer intervals where sessions were terminated after 3 hours. A stepper distributed responses to one of nine counters as a function of the time since the last delivery of food.

In a second experiment (1b) four naive pigeons experienced interfood intervals of 5, 10, 15, and 30 sec, according to a Latin square design. For these animals no DRO contingency was employed.

Results

The average response rate for each animal in each condition may be found in Table 1. Figure 1 shows the rate of activity as a function of the time since the last delivery of food. The data are averages (geometric means) over four animals, averaged over the last five sessions at each condition of Experiment 1a. The abscissa has been normalized for convenience in presenting the data.

The data for the three shorter interfood intervals all show a pronounced increase in activity after feeding, an increase that is maximal one-fifth of the way through the interval. The overall rate of activity is greatest for the shortest interfood intervals and decreases monotonically with increasing interfood intervals.

Response rates were substantially higher in Experiment 1b, but the time-course of activity through the interfood interval was similar. Activity was maximal one-quarter of the way into the interval and was minimal immediately after feeding. The decrease in activity before feeding was not nearly so marked as in Experiment 1a, presumably because no DRO contingency was employed.

TABLE 1

INTERFOOD INTERVAL, RESPONSE RATE, AND
PARAMETER VALUES FOR EQUATION 1 IN
EXPERIMENTS 1a AND 1b

Interfood interval (sec)	Re- sponses per minute	A	C	I	n ^a
Experiment 1b					
5	61.2	600	.37	.24	76
10	63.6	318	.46	.20	80
15	58.4	185	.57	.16	75
30	55.1	130	.86	.16	89
Experiment 1a					
25	21.0	134	.30	.13	24
60	15.6	47	.44	.07	18
120	10.7	32	.46	.07	12
200	7.1	10	1.63	.02	16
400	5.7	7	3.42	.00	22

^a Number of parallel processes necessary to account for the data if they are assumed to derive from a series latency mechanism.

Discussion

A surprising amount of general activity occurred in this feeding situation despite contingencies that reinforced quiescence. Other investigators have also noted the difficulty of getting animals to stand still using food as a reward (Blough, 1958; Breland & Breland, 1966). The situation is the inverse of that dealt with by Bolles (1970b) and Keehn (1972), who noted the ease of getting rats to "freeze" in the context of occasional electric shock. Bolles proposed the existence of various "species-specific defense reactions," such as freezing, fleeing, and fighting, which were especially easy to condition with aversive stimuli. It may well be that the floor panels of the chamber used in these experiments provide a gross measure of "species-specific appetitive reactions," behaviors that are either induced by the recurring food deliveries, or are easily conditioned through adventitious pairing with food. The latter notion is suspect, however, since the 5-sec DRO was specifically designed to minimize superstitious behaviors.

It is curious that in both experiments activity decreased before the delivery of food. This decrease was arranged by the DRO contingency of the first experiment, but in Experiment 1b there was no such contingency. It is possible that the motivation

generated by food is simply exhausted by the end of the interfood interval. This is unlikely, however, for all of the curves peak at the same relative time after food delivery and fall to the same level before the next delivery. Were these changes coincident on an absolute time scale, the exhaustion of a reserve would be a plausible explanation. Their coincidence on a relative time scale denotes a greater dependence on the *schedule* of food delivery than on the aftereffect of any *single* delivery.

A more likely explanation for the decrease in activity is that competing behaviors, incompatible with general activity, are elicited by the temporal proximity of food. These activities might include consistent orientations to the front of the chamber, pecking at the front wall, and investigation of the magazine aperture. Such behaviors have been observed at the end of a free-food interval by Staddon and Simmelhag (1971), and at the end of a conditioned stimulus by Zamble and Kirkwood (1969); they would clearly

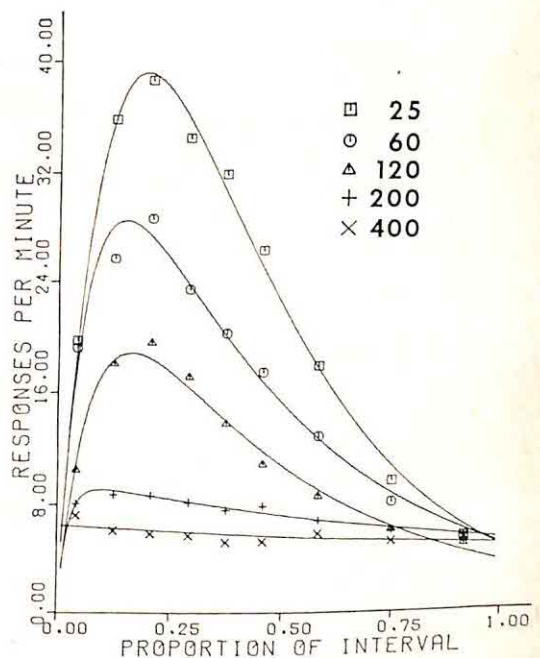


FIGURE 1. Rate of panel activation as a function of the time since reinforcement. (Interfood interval is the parameter; the abscissa is normalized to range between 0 and 1. Data, from Experiment 1a, are averages over four pigeons. The curves are derived from Equation 1.)

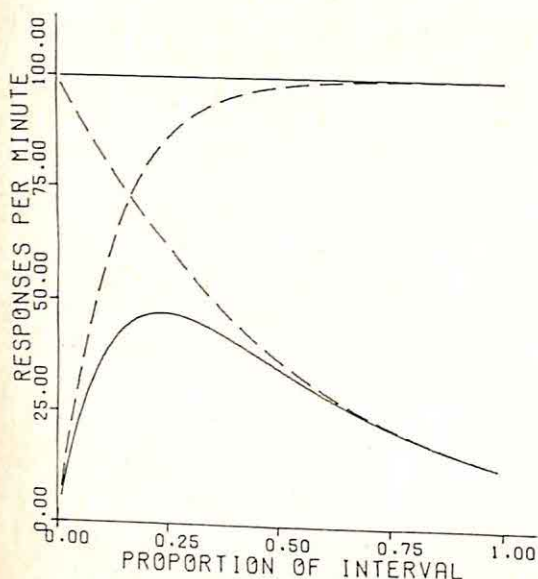


FIGURE 2. Hypothetical curves representing: A constant asymptotic rate of A responses per minute (solid horizontal line), general activity with no competing terminal behaviors (rising dashed curve, from Equation 3), terminal behaviors suffering no inhibition by reinforcement (falling dashed curve, from Equation 4), and the time course of general activity (bitonic solid line, from Equation 1).

interfere with general activity as measured in the present experiment. A more careful analysis of prefood behaviors is conducted in the second series of experiments.

A MATHEMATICAL MODEL

The theoretical curves accompanying the data are derived from the equation

$$R = A(e^{-t/C} - e^{-t/I}). \quad [1]$$

The dependent variable, R , is rate in responses per minute, and e is the base of the natural logarithms. Values for the parameters A , C , and I , which were derived by an iterative least-squares program, may be found in Table 1. The proportion of the interfood interval measured by t ranges between 0 and 1.0. The parameter A is a scale factor, affected by the units of measurement and the sensitivity of the system. Of greater theoretical importance, A provides a measure of the maximum level of activity engendered by the feeding schedules. In the absence of competing behaviors or inhibitory stimuli, Equation 1 would predict a constant rate of A panel activations

per second (cf. top solid line, Figure 2). If from that constant rate we subtract

$$R = Ae^{-t/I}, \quad [2]$$

we obtain

$$R = A(1 - e^{-t/I}), \quad [3]$$

whose locus is traced by the rising dashed line in Figure 2. Parameter I governs the rate at which the response curve approaches maximum (A); the smaller I , the steeper the ascending part of the curve. Parameter I is interpreted as a measure of the inhibition of activity by food delivery which is believed due to the role of food in signaling a period during which the next food delivery is unlikely (cf. Jensen & Fallon, 1973; Siegel & Domjan, 1971; Staddon, 1974).

Consider now the equation

$$R = Ae^{-t/C}, \quad [4]$$

which is traced as the falling dashed line in Figure 2. This curve accurately represents the right tail of the activity curve, approaching 0 as t grows large. The decrease in activity before the delivery of food (at $t = 1.0$) is presumed to be due to the presence of competing behaviors which occur toward the end of the interfood interval and which interfere with general activity. In Experiment 2, the rate of some terminal behaviors is modified directly by reinforcement, and the predicted changes in C are witnessed.

When these two basic processes are joined—the fast acting inhibition of Equation 2 being subtracted from the slower decay of general activity modeled by Equation 4—Equation 1 and its curve at the bottom of Figure 2 result. Equation 1, which was empirically derived, has close relatives in at least two other formal systems, control theory and stochastic latency models.

Control, or cybernetic, systems are those which employ some form of feedback to moderate the output of the system (cf. McFarland, 1971; Milsum, 1966). Examples range from servo-mechanisms to thermostats, and in humans include the gamma-efferent control of limb movement and the mechanisms which govern pupil aperture. Random impulses to such homeostatic systems will cause momentary disruption, followed by asymptotic return to the set-point.

One of the simplest models of this asymptotic return is provided by the exponential function. The impulse carries the system to some level A , and this is followed by a return to the reference level whose course is described by an equation such as 4. Immediately after the impulse ($t = 0$) the system is at Level A , and as t grows large the function approaches 0. The rate of return to reference level (0) is governed by C , which is called the "time-constant" of the system. When $t = C$, the system has completed 63% of its return to equilibrium.

Equation 1 models a more complicated system in which the disruption is not maximal until some time after the impulse. An analogue is found in the striking of a very elastic ball. The energy of the impulse is first absorbed in deforming the ball, and only moments later is effective in propelling the ball through space. Equation 1 represents a second-order control system with two time constants, C and I . If the restoring force, $1/C$, were 0 (cf. Equation 3), the system would come to rest at Level A and would be 63% of the way to that level after a time equal to I . For C s less than infinity, one contends with the difference between two exponential curves, one tending to carry the system asymptotically to Level A and the other tending to carry it asymptotically back to the reference level. The form of the resultant curves depends on the values of the time constants, with those found in Figure 1 being typical examples.

The reciprocal of the geometric mean of C and I defines the "resonant," or "natural" frequency (ω) of a control system. This index remained approximately constant at 5.4 in Experiment 1a and at 3.2 in Experiment 1b. The value of ω reflects the responsiveness of the system; since the obtained values are "faster" than the input frequency (i.e., the rate of feeding, here given a value of 1), the level of activity should rise to its maximum and drop back to equilibrium relatively quickly. In the normalized coordinates of Figure 1, the constancy of ω indicates that the natural frequency of the system is a linear function of the actual interfood interval—the organism "stays in-tune" with its feeding schedule. The cybernetic interpretation

of Equation 1 will be further examined later in the General Discussion section.

Equation 1 is also found in the analysis of stochastic latency mechanisms (McGill, 1963), as one instance of a "general-gamma distribution" (McGill & Gibbon, 1965). If an event such as "Response A" has a fixed probability of occurring in any small unit of time, then the probability that the response will occur by time t , $f(t)$ has a geometric distribution that is approximated by the exponential:

$$f(t) = \lambda e^{-\lambda t}. \quad [5]$$

Consider now a series latency mechanism wherein the response latency has two components,

$$t = t_1 + t_2. \quad [6]$$

The values t_1 and t_2 are random delays associated with the two components which must be completed in sequence before a response is recorded, and

$$f(t_1) = e^{-t_1/I}/I, \quad [7]$$

$$f(t_2) = e^{-t_2/C}/C. \quad [8]$$

Then

$$f(t) = \frac{1}{C - I} (e^{-t/C} - e^{-t/I}). \quad [9]$$

Equation 9 resembles Equation 1, although each describes a quite different set of data—the former changes in response rate, the latter a probability density function. The curves may be made congruent by assuming that each panel activation is the outcome of a unique two-component series latency mechanism (Equation 9) and that many such mechanisms operate in parallel. The random completion of each of the two-step processes throughout the interval will result in a single measured response; the curves of Figure 1 result from the completion of many such processes. The first event in this two-step process, specified by Equation 7, need not be an observable response or any sort of response as those are normally construed. It may be conceived as the timing-out of an internal clock which signals the end of a quiescent period. There ensues a second period of time terminated by a recorded response.

One of the quick returns on an investment in stochastic latency mechanisms is an esti-

mate of the number of parallel processes necessary to account for the data. Since each process will generate a function with a coefficient of $1/(C - I)$, n processes are necessary to generate a function with a coefficient of A , where $n = A(C - I)$. Values of n are listed in Table 1. It can be seen that the estimates of n range around 80 for Experiment 1b and around 18 for Experiment 1a. Because A incorporates a scale factor, the estimates are unique up to a multiplicative constant. The estimates show no dependence on interfood interval but do show a large increase in Experiment 1b, where activity was no longer suppressed by a DRO contingency. The correlation of the hypothetical gamma processes with observed behaviors and their employment as minimal units of behavior are the types of potential benefits which might justify the assumptions of underlying series latency mechanisms.

If an additional step is added to the series latency mechanism, Equation 10 results:

$$R = N \left(\frac{Ce^{-t/C}}{(C - I)(C - B)} + \frac{Ie^{-t/I}}{(I - C)(I - B)} + \frac{Be^{-t/B}}{(B - C)(B - I)} \right) \quad [10]$$

It can be seen that as B (the time constant of the third step) approaches zero, Equation 10 reduces to Equation 9. In the system modeled by Equation 10, each recorded response is the outcome of a three-step chain. This equation is especially useful as a model for behaviors occurring during brief periods of time—interfood intervals or conditioned stimulus durations of less than 15 sec. The consumption of grain or the orientation to the conditioned stimulus may extend several seconds into the interval; where longer intervals are under consideration, the error attributable to such spillover behaviors is negligible. For shorter intervals it becomes worth incorporating an additional step (and parameter) into the model. Whereas Equation 1 accounted for 95% of the data variance of the two shortest intervals in Experiment 1b, Equation 10 accounted for over 98.5% of that variance.

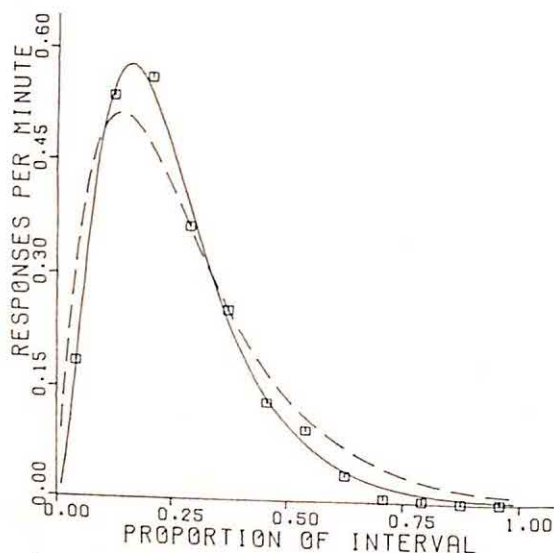


FIGURE 3. The probability of pigeons emitting some interim behavior. (Based on data from Staddon & Simmelhag, 1971; the dashed curve was derived from Equation 1, the solid curve from Equation 10.)

Additional data for which Equation 10 is a useful model were collected by Staddon and Simmelhag (1971), who monitored the activities of pigeons that were given access to grain every 12 sec. Staddon and Simmelhag dichotomized those activities into a set occurring just before food delivery ("terminal behaviors") and a set consisting of all other activities ("interim behaviors"). I have combined their data to reflect the probability that any interim behavior occurs during each second of the interval. These data are displayed in Figure 3, along with curves derived from Equation 1 (dashed) and Equation 10 (solid). Although the value of B was quite small (.05, implying that 63% of the spillover behaviors were completed within the first .6 sec), it improved the proportion of variance accounted for from .935 to .995.

As additional steps are added to the series latency mechanism, the probability density functions quickly become unmanageable. At the same time, however, those functions converge on the normal distribution:

$$R = He^{-\left(\frac{t-M}{S}\right)^2} \quad [11]$$

where M is the midpoint of the curve, S is proportional to the standard deviation ($S = \sqrt{2} \sigma$), and H is the height of the curve at

$t = M$. Equation 11 will subsequently be used to describe the time course of various adjunctive and terminal behaviors.

CONCURRENT AND COMPETING OPERANT BEHAVIORS

The parameter C has been identified as descriptive of behavior occurring predominantly near the end of the interval. The smaller C , the more quickly activity subsides to near-zero levels. In the first experiment, a concurrent operant response—key pecking—is established. Insofar as this response is incompatible with general activity, a further suppression of responding should occur near the end of the interval and should be reflected in a decrease in C . In the second experiment, general activity is first suppressed with increasingly severe DRO contingencies and then is reinforced with a positive contingency between activity and feeding.

Experiment 2a

Two pigeons were trained to peck at a Gerbrands response key centrally located above the hopper in the experimental chamber, the first key peck to occur 60 sec after the last reinforcement activated the hopper for 3 sec.

The introduction of a contingency requiring a key-peck response appeared to have had little effect on the pattern of general activity, although the overall level of responding was lower than that obtained for the same animals during Experiment 1a. But examination of the parameters leads to a different conclusion. A and I did not change from control values (.54 and .09) when the concurrent response was introduced,¹ but C decreased from .38 to .24. The lowering of activity level was attributable solely to a change of the parameter governing the rate of return to equilibrium.

Experiment 2b

Three naive pigeons were maintained at 80% of their free-feeding weights. Three seconds of access to grain, normally occur-

ring every 60 sec, could be delayed by panel activations occurring during the last few seconds of the interval: DRO contingencies of 0, 5, and 12 sec were employed in ascending order.

In the last condition of this experiment, the DRO was replaced with a fixed-ratio schedule. After 55 sec had elapsed, each panel activation would count toward reinforcement, which was delivered after 10 activation responses. The purpose of this scheduling was to encourage high response rates while attempting to maintain constant the interfood interval.

Results. For all conditions except fixed ratio, rate was greatest about one-fifth of the way through the interval. Overall rate of responding decreased with increases in DRO length (cf. Table 2). The presence of a DRO contingency affected the shape of the curves, but even without that contingency, activity decreased before reinforcement. The fixed-ratio contingency increased response rates, but the shape of the activity curve was not the familiar "scallop" or the "break and run" usually obtained on ratio schedules. Instead, activity was greatest about halfway through the interval and decreased slightly before reinforcement. These data are essentially identical to those collected in a pilot study with experienced subjects.

Once again the data have been fit with curves derived from Equation 1, with parameters listed in Table 2. The radical changes in terminal contingencies have no effect on I . Changes in the terminal contingencies are correlated with changes in C , which de-

TABLE 2
EXPERIMENTAL CONDITION, OBTAINED INTERFOOD INTERVAL, AND PARAMETERS OF EQUATION 1 FOR EXPERIMENT 2b

Terminal contingency	Interfood interval (secs)	A	C	I
DRO 0	60	39.7	.87	.05
DRO 5	67	36.6	.68	.05
DRO 12	75	24.1	.52	.05
Fixed ratio 10	98	22.5	7.14	.05

Note. Abbreviation: DRO = Differential Reinforcement of Other Behavior.

¹ Parameters were said to remain invariant if, when left free to vary, the data variance accounted for by the model increased by less than 1%.

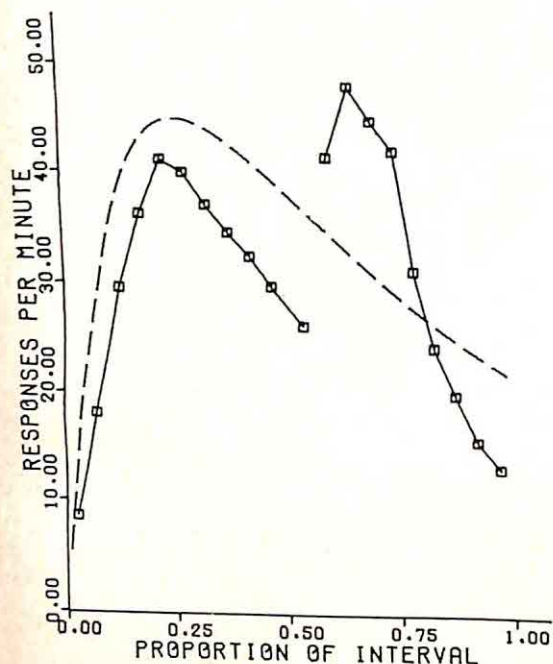


FIGURE 4. Rate of panel activation in Experiment 3a before blackout (.0 to .5) and before reinforcement (.5 to 1.0). (The dashed line represents the locus of the control data.)

creased with increasing DRO requirements, and increased markedly under the fixed-ratio contingency.

In Experiment 2a, although the level of general activity decreased upon introduction of the competing behavior, the theoretical asymptotic level of activity (A) did not change. In the present experiment, the value of A did vary with changes in the terminal contingencies, although the values were not monotonically related to the changes in terminal response rates. It is probable that changes in A were due solely to variations in the obtained interval, for imposition of the DRO contingencies decreased the rate of reinforcement (see Table 2), and a further decrease occurred when the fixed-ratio contingency was introduced.

Discussion. These experiments show that activity may be suppressed, either by the introduction of competing behaviors or by direct punishment of that activity. Equation 1 remains an appropriate model, with the effects of these manipulations primarily manifested in one parameter, C.

EFFECTS OF INTERPOLATED STIMULI

The curves in Figure 1 are reminiscent of curves drawn by Konorski (1967) to depict the "hunger drive" which occurs in anticipation of food, is suppressed during eating, and is "restored with rebound after the [eating] response is over" (p. 46). Their nature is more complex, however, than the simple model proposed by Konorski. The peak "rebound" of these curves does not bear a fixed temporal relationship to the eating response: Instead, the maximum occurs a fixed proportion of the distance until the next feeding. This anticipatory nature of the data seems to rule out a simple reflexive interpretation, such as Konorski's.

In the first series of experiments two effects of reinforcement were thoroughly confounded. The delivery of food was both the main source of energy or motivation for the subjects, and thereby undoubtedly had direct effects on their behavior, and it was also a signal with a fixed temporal relationship to the subsequent reinforcement, and thereby undoubtedly acquired incentive motivational (Bindra, 1968) or inhibitory (Staddon & Innis, 1969; Staddon, 1970) properties. In the following series of experiments additional stimuli are added in various temporal relations to reinforcement, and are found to have patent inhibitory and excitatory effects upon the course of activity.

Experiment 3a

Method. Four pigeons were exposed to a schedule of free food every 60 sec. After nine sessions, every other feeding was replaced with a 3 sec blackout of the houselight, and an additional nine sessions were conducted. Activity in the chamber had no effect on the scheduling of blackout or food.

Results. Response rates before blackout (0–50% of interval) and after blackout (50–100% of interval) are shown in Figure 4, along with the locus of response rates during the control condition (dashed line). Responding after feeding shows an increase to maximum rate about one quarter of the way through the interval, followed by a slower decrease until the blackout occurs. Activity after blackout begins substantially below the maximum postreinforcement rate, then in-

creases quite dramatically, peaking at a rate higher than that found after food, and even higher than the maximum rate during 60 sec interval pretraining of these subjects.

The present data seem consistent with studies of the "Frustration Effect" (Amsel, 1962) or of the patterning of running in studies of alternate reinforcement in straight alleys (Capaldi, 1966, 1970). In both cases, running speed is greater on the trials after nonreinforcement than after reinforcement. Indeed, the present curves might easily be taken for running-speed data, for they show the characteristic acceleration and deceleration associated with the movement of a rat in the confines of an alley (Amsel, 1972).

The decrease in rate before reinforcement is similar to that found in all of the experiments reported thus far, and, as in those cases, is most likely due to terminal behaviors such as "hopper-tending" that are elicited by the temporal proximity to reinforcement. The decrease in rate before blackout is a more curious affair. Blackout will occur no matter where the animal is in the chamber, and there are no behaviors which might possibly increase the value of that stimulus change, as hopper-tending might increase the amount of food the animals can obtain. Since pigeons tend to "roost" in the dark, it is possible that roosting behavior was conditioned to occur during the preblackout period. A more likely explanation is that the two intervals were confused by the pigeons, and the prereinforcement behaviors generalized to the preblackout period. These hypotheses are examined in the following two experiments.

Experiment 3b

The decrease in responding before blackout in Experiment 3a might have been due to the generalization of prereinforcement behaviors into the preblackout period. To test this possibility, the two periods were made quite discriminable: the houselight was turned off after 60 sec had elapsed since reinforcement and was left off until the end of the next reinforcement. Thus, during the first half of the 120-sec interfood interval the chamber was normally illuminated, and during the second half it was blacked out. The

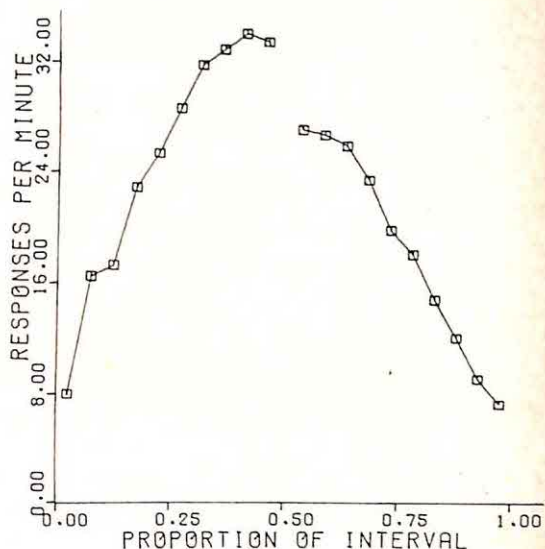


FIGURE 5. Rate of panel activation in Experiment 3b before blackout (.0 to .5) and during blackout (.5 to 1.0).

subjects and apparatus were the same as in the previous experiment.

Results. Figure 5 shows that the clear differentiation of the intervals eliminated the downturn in the postreinforcement activity curve: The animals increased their activity up until the blackout (at 50% of the interval) and then decreased it until the delivery of reinforcement. Note furthermore that the pigeons were not stationary during the blackout, and that this much longer blackout did not elicit preblackout quiescence.

Experiment 3c

In the preceding experiments, blackouts interpolated within the interfood interval appeared to have various facilitative and inhibitory effects on the behaviors preceding and following them. Some of the effects were demonstrated to due to trace-discriminative control or to the generalization of prereinforcement behaviors to the prestimulus period. Whereas there appeared to be an abundance of general activity in these experiments and a facilitation of activity by an interpolated blackout (in Experiment 3a), general activity has always decreased before food delivery. This seems inconsistent with observations that

incentive-motivational stimuli are sufficient to increase general activity . . . [this] activity has all

the marks of anticipatory excitement, such as that seen during classical appetitive conditioning . . . and that described by ethologists as 'appetitive behaviors'. (Bindra, 1968, pp. 3-4)

The present experiment was conducted in an attempt to more clearly demonstrate "anticipatory excitement." A flickering houselight was used to signal the imminent delivery of food. This stimulus was less likely to elicit roosting, was long enough to permit a measure of activity during its occurrence, and was more closely associated with food delivery than were any of the stimuli used in the preceding experiments.

Method. Two of the pigeons from the previous experiments were given 3-sec access to mixed grain every 60 sec. After seven sessions of these control conditions, the basic feeding schedule was continued for an additional seven sessions, but during the last 12 sec of each interval the houselight was pulsed at a rate of 10 Hz.

Results and discussion. Figure 6 shows the experimental (squares and solid line) and control (circles and dashed line) data. The 12-sec conditioned stimulus inhibited activity immediately before reinforcement, but released from inhibition behaviors occurring

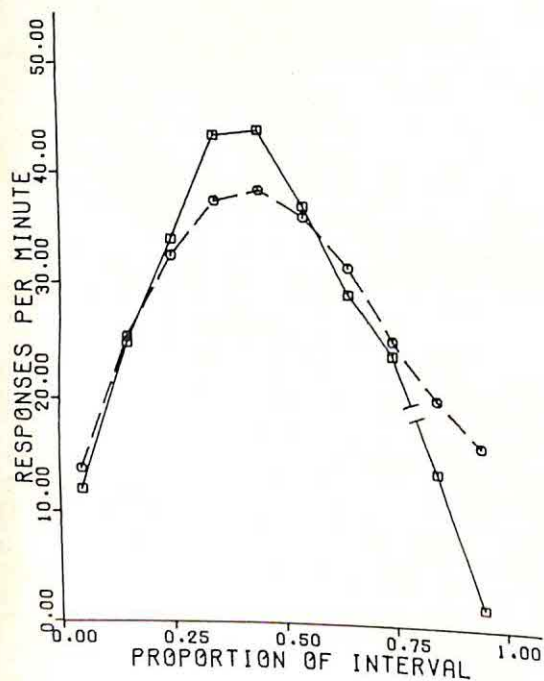


FIGURE 6. Rate of panel activation in Experiment 3c under control conditions (dashed line) and experimental conditions (solid line). The CS began at 0.8.

earlier in the interval. These somewhat paradoxical effects may be understood in terms formulated by Staddon and Simmelhag (1971) in their analysis of "superstitious" behaviors.

Staddon and Simmelhag monitored the behavior of pigeons who were given 3-sec access to food every 22 sec (cf. Experiment 1b of this review). They noted many forms of activity during the interfood interval, some of which followed time courses similar to those found in the present experiment and some of which started later in the interval and increased in rate as the moment of reinforcement neared. The former behavior they termed "interim" responses, the latter "terminal" responses. Staddon and Simmelhag suggested that interim behaviors reflect a set of general activities instigated directly by reward. It is from this set of activities that a few appropriate behavior patterns may later be selected, via a mechanism such as the law of effect. Reinforcement thus has two aspects, "variation"—the generation of diffuse and generalized interim activities—and "selection"—the strengthening of particular behaviors selected by contiguity to reinforcement. Staddon and Simmelhag's analysis and evolutionary model are fundamental to the behavior studied in this article. An analogue may be drawn between the rise of interim behaviors and the parameter I and between their displacement by terminal behaviors and the parameter C.

What terminal behaviors might be expected of a pigeon? Pecking, hopper observation, and other behaviors that have either an innate or learned association with the delivery of food. These behaviors all leave the animal standing in one spot and therefore compete with the explicit dependent variable in the present experiments, general activity. In Experiment 3c, the conditioned stimulus was a less ambiguous predictor of the imminence of reinforcement than whatever temporal cues might have been available to the pigeons, and thus elicited, or was an S^D for, terminal activity. At the same time, the absence of the conditioned stimulus became a predictor of non-reinforcement, and by inhibiting terminal behaviors, it disinhibited interim activity. This

analysis is not inconsistent with Bindra's (1968) theory of incentive-motivation. The terminal behaviors of dogs and rats, Bindra's reference subjects (Bindra & Palfai, 1967; Zener, 1937), were likely to be vigorous coursing or foraging responses that would contribute to measures of general activity. These and subsequent experiments lend confirmation to Bindra's later (1970, 1972) development of incentive-motivation theory.

General activity thus appears too gross a measure to be useful in comparing levels of motivation across species. Indeed, it will not do even for a single species: Levels of activity immediately before reinforcement in Experiment 1a were the same for the 400-sec and 25-sec conditions, but, by most accounts, animals in the former condition would be judged as more highly aroused. This presumed higher level of arousal is reflected in a greater value of A for the shorter interval. Furthermore, while there was little difference in response rates on the 5-, 10-, and 15-sec schedules of Experiment 1b, A continued to show a proportional increase with increases in rate of feeding ($A = 3,000/T$, where T is the interfood interval), and so provides an index of "arousal" that is unbiased by "ceiling effects" on obtained response rates. It is possible, therefore, that the theoretical asymptote for activity, A , will provide a useful reification of the construct arousal within a species, and, with recalibration, may be of value in cross-species comparisons.

Experiment 3d

By providing a more accurate temporal reference point, signals of forthcoming reinforcement conduce a more intense surge of either interim or terminal behaviors. In Experiment 3a, the blackout permitted a brief disinhibition of activity followed by a more intense suppression. In Experiment 3c, the effects were primarily inhibitory, though a finer analysis of behavior during the first few seconds of the condition stimulus might have uncovered some excitation. This last effect is somewhat curious, since in Experiment 1b, intervals of 5 and 10 sec showed marked patterns of excitation. Judging from those conditions, there was ample time for the

animals to move about the box during the 12-sec conditioned stimulus (CS) in Experiment 3c. Was inhibition predominant in this experiment because the animals had already had an opportunity, during the pre-CS period, to "displace" their energies? Or is the important variable in determining the predominance of interim or terminal behaviors the relative, rather than the absolute, proximity to reinforcement (cf. Jenkins, 1970; Staddon, 1972)? In Experiment 1 the relative proximity of the signal (reinforcement) to reinforcement was zero. Reinforcement was the event most distant from the next reinforcement and thus inhibited the occurrence of both interim and terminal behaviors. In Experiment 3a the relative proximity of the signal (blackout) was .5, and both interim and terminal behaviors were intensified. In Experiment 3c, relative proximity was .8, and only terminal behaviors were observed. In the following experiment, the relative proximity will be set at .8, but the interval will be lengthened to 300 sec. Will a signal with the same relative proximity to reward, but with a much smaller absolute proximity, still incite only terminal behaviors?

Method. The subjects had served in Experiment 1b. Food delivery was scheduled to occur every 300 sec and was independent of the behavior of the pigeons. The houselight was on and was constant until 60 sec before food delivery, at which time it was pulsed at a frequency of 10 Hz. After 10 sessions, 10 control sessions were conducted in which food was delivered every 60 sec with the houselight constantly on. Finally, 10 additional sessions were conducted during which the houselight was off until 60 sec before food delivery and then was turned on and pulsed.

Results and discussion. In Experiment 3c, the houselight was pulsed during the last 20% of the interval and during that time general activity was minimal. In the present experiment the houselight was on during the last 20% of the interval, during which time a familiar bitonic trend of general activity through the interval was observed (see Figure 7, squares). An average of 15 panel activations occurred during the 60-sec conditioned stimulus, half of the rate under the control condition (dashed line). While the response rate during the conditioned stimulus

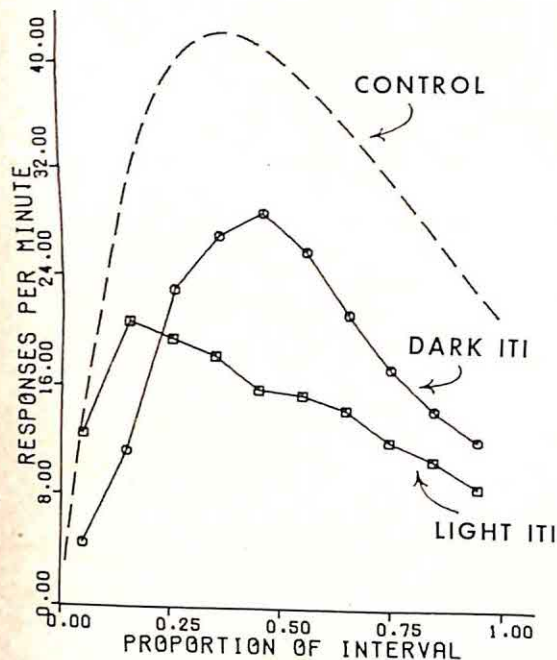


FIGURE 7. Rate of panel activation in Experiment 3d under control conditions (dashed line), first experimental condition (houselight on - pulsed; light ITI, squares), and second experimental condition (houselight off - pulsed; dark ITI, circles).

was greater than the rate immediately preceding the conditioned stimulus, it was lower than the average rate of activity during the 240 sec preceding CS onset (21 responses per minute).

When the houselight was turned off during the intertrial interval, responding was completely suppressed until the onset of the conditioned stimulus, and then it increased to a peak of 28 responses per minute approximately halfway through the CS period (Figure 7, circles). The shapes of the curves differed, with post blackout responding showing the slowest rise to its peak.

Discussion

Relatively more activity was elicited by the conditioned stimulus in Experiment 3d than was the case in Experiment 3c. While the conditioned stimulus elicited more activity than might otherwise be expected during this segment of a 300-sec interfood interval, rate did not exceed the average rate throughout the rest of the interval. A more dramatic effect was obtained when the chamber was darkened during the 240-sec intertrial in-

terval, whereupon rate during CS increased substantially.

Evaluation of the present data in terms of average rates would leave the most important aspects unseen. A signal of forthcoming reinforcement is neither wholly inhibitory nor excitatory but has the now familiar multiphasic effect on activity. Whether the signal is the shift to a flickering houselight or the delivery of food, immediately thereafter activity is suppressed. This suppression occurs whether the conditioned stimulus is of 12-sec or 60-sec duration. An increase in activity may occur after the inhibition; such an increase was noticeable for the 60-sec CS—especially when the conditioned stimulus was the transition from houselight-off to houselight-on—but might have occurred unmeasured for the shorter duration CS. In all cases there was a decrease in activity immediately before reinforcement, presumably due to the evocation of competing terminal activities.

The present series of experiments demonstrate that the course of general activity is strongly affected by the presence and nature of signals within the interfood interval. General activity is not a simple reflexive response to food delivery, for it is modified by signals of impending food—inhibited by signals with close proximity to food, disinhibited and then inhibited more intensely by signals with moderate proximity to food. The last experiment generated activity curves during a conditioned stimulus that resembled the curves obtained in Experiment 1 where the only conditioned stimulus was food delivery. An adequate discussion of these effects will require a vocabulary that includes "words" such as Equation 1, for verbal description of such phenomena will otherwise be ponderous and will omit important information, such as the time constants of the component processes.

DRUG EFFECTS

The parameters of Equation 1 have a relatively straightforward interpretation and summarize effectively various aspects of performance. In the following experiments, the robustness of the basic model is tested. Pigeons will be subjected to various drugs

and deprivation levels. To the extent that these perturbations have selective effects on response systems as modeled by Equation 1, they will be reflected in selective changes in the parameters of that model. To the extent that new processes are introduced by the experimental operation, Equation 1 will fail.

Method. Four White Carneaux pigeons served as subjects. They received 3-sec access to mixed grain every 60 sec; this schedule was independent of any behavior on their part. They were exposed to the following conditions in the order listed. Sessions were conducted daily, with two administrations of each drug, one on Wednesday and one on Friday. Control data are from the Tuesdays preceding drug administrations and from the days preceding the two prefeeding sessions.

Prefeeding. Each pigeon was given 60 grams of grain 1 hour prior to its normal session time. This amount equalled 12% of their experimental weight. After the first prefeeding session, the animals returned to their 80% weight, and additional control sessions were conducted before the second prefeeding session.

Caffeine. Each pigeon was given 10 mg of caffeine (approximately 20 mg/kg) 1 hour prior to its normal session time. Route of administration was oral.

Acepromazine maleate. Each pigeon was given 12 mg/kg of the common animal tranquilizer acepromazine maleate (Anatran) 1 hour prior to its normal session time. Route of administration was oral.

Scopolamine. Each pigeon was given .5 mg/kg of scopolamine 15 minutes prior to its normal session time. The scopolamine was injected intraperitoneally.

Results and Discussion

The effects of the experimental manipulations on response rates were generally in the expected directions. Prefeeding greatly decreased rate, while caffeine increased rate. Both the tranquilizer Anatran and scopolamine brought about slight decreases in rate (see Table 3). Inspection of Figure 8 confirms these effects. There also appears to be some change in the shapes of the curves, which may be analyzed by inspection of the parameters.

Prefeeding the animals, which brought about a fourfold decrease in rate, left the shape of the curve unchanged from control conditions: While A decreased from 74 to 19, C and I remained unchanged. This invariance is important, for it suggests that decreases in motivation due to satiation, as

TABLE 3
RESPONSE RATE AND PARAMETERS OF
EQUATION 1 FOR EXPERIMENT 4

Condition	Re- sponses per minute	A	C	I
Control	48.0	73.9	1.61	.10
Prefeeding	12.9	19.0	1.61	.10
Caffeine	51.3	149.0	.74	.21
Acepromazine maleate	40.5	50.7	2.81	.06
Scopolamine	45.8	62.1	2.63	.10

might occur within the course of an experimental session, will have a linear effect on performance at all points within the interval. Furthermore, differential effects of drugs on C and I cannot be explained as a simple appetite effect, for a substantial prefeeding generates no such effects.

The large dose of caffeine brought about only a small increase in rate of panel activation. Inspection of Table 3 and Figure 8 shows, however, that there were large effects

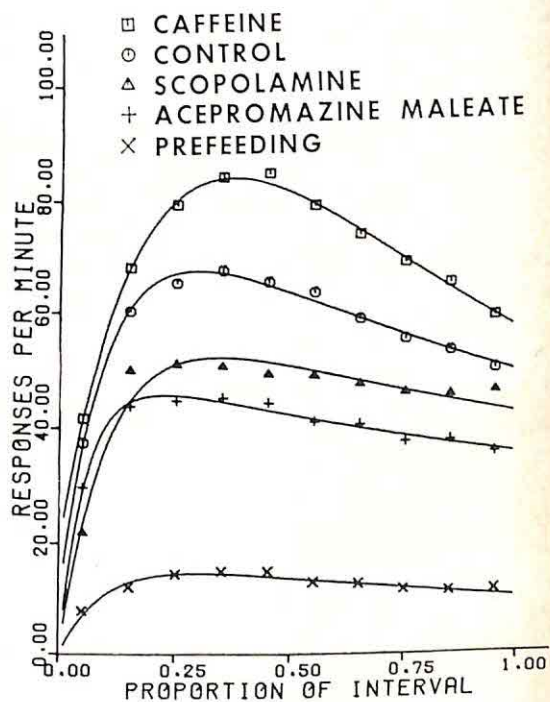


FIGURE 8. Rate of panel activation in Experiment 4b. (For clarity of presentation, the top curve, caffeine, has been displaced upward by 20 responses per minute, and the second curve, control, has been displaced upward by 10 responses per minute.)

on behavior that are not reflected in an average rate measure. The asymptotic rate, A , doubled, while the rate of approach to asymptote decreased. The activity then subsided at a faster rate than under the other conditions of this experiment, with C decreasing to less than half of the control value.

The changes under caffeine may be interpreted as an intensification both of general activity and of the terminal behaviors which compete with general activity toward the end of the interval. The slower rate of approach to asymptote might be interpreted as an intensification of "initial" behaviors which compete with interim behaviors at the beginning of the interval. But there were no obvious initial behaviors apparent during casual observations of the pigeons under various conditions of these experiments. The resonant frequency under caffeine, $1/\sqrt{CI} = 2.52$, was close to that of the control condition, 2.49.

The effects of acepromazine maleate were opposite to those of caffeine. Asymptote decreased, rate of approach to asymptote increased, and rate of recovery decreased. Resonant frequency remained close to the control value, at 2.53.

The effects of scopolamine were curious. There was a slight decrease both in average response rate and in asymptote. The rate of approach to asymptote remained at the control value, while the rate of recovery decreased. This asymmetric change in the parameters caused the resonant frequency to decrease to 1.95.

Scopolamine is known to cause a decrease in habituation to novel stimuli and a failure to inhibit incorrect responses in learning situations (cf. Carlton, 1969). The increase in C reflects a failure to inhibit general activity at a time when terminal behaviors might be more appropriate. The lack of effect on I indicates that scopolamine has no effect on the excitation of general activity but only on the excitation of specific terminal behaviors which compete with general activity.

The behavioral changes brought about in this experiment are consistent with current intuitions concerning the nature of the experimental manipulations, and thus they rein-

force both those intuitions and the use of the model for their quantitative assessment. The dissociation of changes in C and I effected by scopolamine may reflect different neurochemical bases for induced (interim) and operant (terminal) behaviors (cf. Margules & Margules, 1973). The invariances that were found—the constancy of the parameters under prefeeding and the constancy of the resonant frequency under all but the scopolamine conditions—challenge the various interpretations of the model. These issues will be examined later in the General Discussion.

ADJUNCTIVE BEHAVIORS

Experiment 5a

In training animals to perform for vaudeville shows, Keller and Marion Breland (1961) noted that the form of trained operant behaviors would often drift toward that of more species-typical instinctive behaviors and that instinctive behaviors would intrude into operant chains. A prototypical example may be found in sideshows today when one commands a chicken to "play a piano" by inserting a quarter into its "modified Skinner box." The performance is frequently interrupted by vigorous scratching of the floor of the box, much as if the animal were scratching for seeds in loose dirt. This behavior was analyzed in the following experiment.

Method. Four bantam chickens hens were procured for the experiment. The chickens were reduced to 85% of their free-feeding weight and then were magazine trained using the same mixed grain that was employed for pigeons. The chickens were exposed to a 60-sec feeding schedule with no DRO contingency. Instances of scratching were recorded by depressing a thumbswitch, which incremented the data bins every .1 sec.

Results. All chickens spent a substantial amount of time scratching the floor and front wall of the chamber. Many of the scratches were directed at the hopper aperture. Figure 9 shows the average number of seconds spent scratching as a function of the time since the last feeding. The curve for panel activation was quite similar and is not shown here.

The unconditioned scratching response of chickens shows a patterning throughout the interfood interval not unlike that of the general activity of pigeons. Scratching increased

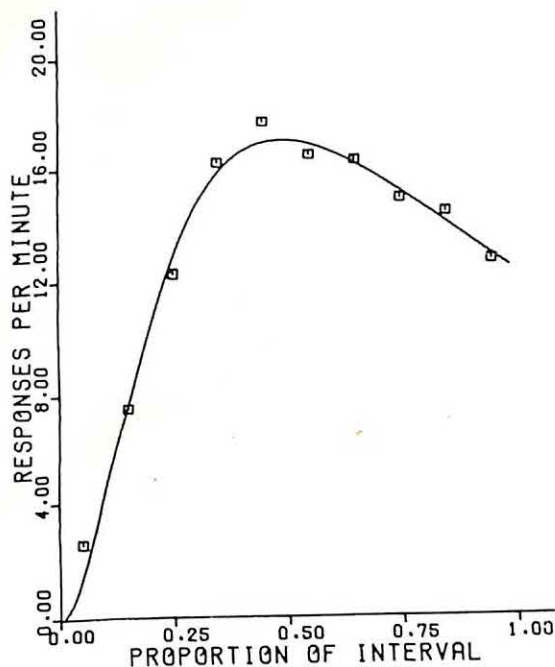


FIGURE 9. Rate of scratching (seconds per minute) induced by reinforcement in Experiment 5a. (The curve was derived from Equation 10.)

to a maximum about halfway through the interval and decreased thereafter. Equation 1 accounted for 95% of the variance in the data, while Equation 10, which is represented in Figure 9, accounted for 98.6%.

Experiment 5b

Another set of behaviors induced by intermittent feeding has been labeled "schedule-induced aggression" (Azrin, Hutchinson, & Hake, 1966; Flory, 1969a). This type of adjunctive behavior was studied in the following experiment.

Method. Hungry pigeons were placed in the chamber, each with an undeprived target pigeon. The experimental and target pigeons were separated by a clear plastic partition 15 cm high, placed diagonally in the chamber. The experimental subjects were fed every 60 sec, independently of their behavior. Pecks directed at the target bird were recorded by depressing a thumb-switch, which incremented the data bins every .1 sec.

Results and discussion. During the last five sessions, one pigeon spent 9 sec out of every minute attacking the target pigeon, while the other spent $5\frac{1}{2}$ sec per minute in attack. Equation 1 was fit to the first seven

data points and accounted for 97.7% of their variance (see Figure 10).

Schedule-induced aggression follows a time course similar to general activity (control data from Experiment 1a for these two pigeons appears as a dashed line in Figure 10), both behaviors sustaining representation as the sum of two exponential functions. Just as general activity is suppressed by the proximity to reinforcement, so too is aggression. There was, however, a self-perpetuating aspect to the aggressive behavior, with animals often becoming "locked into" a battle, and in some cases missing food. This perseveration is evident in Figure 10, where aggression during the last 15 sec of the interval is seen to occur at a rate higher than predicted by the model. A less reactive target, such as a contact-sensitive mirror (Cohen & Looney, 1973) might provide a better preparation when the transactional aspects of behavior must be minimized.

Both schedule-induced and extinction-induced aggression are manifestations of normally occurring behaviors whose probability has been greatly increased by a free-food schedule. While such scheduling is maintained, the temporal pattern of aggression is

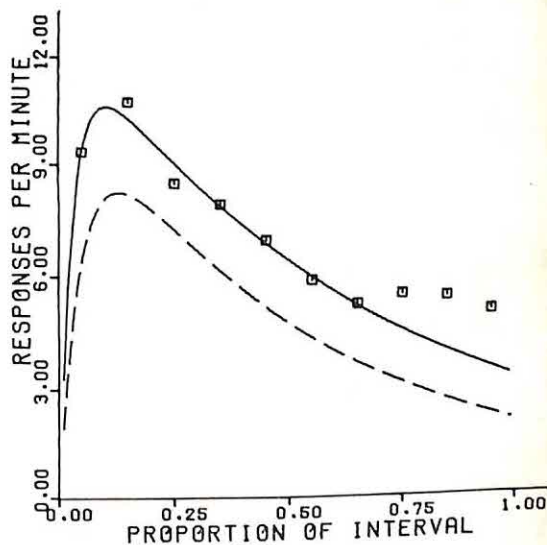


FIGURE 10. Rate of attack (seconds per minute) induced by reinforcement in Experiment 5b. (Also shown is the rate of panel activation for these two pigeons in Experiment 1a, dashed line; scale reduced by .75. The curves were derived from Equation 1.)

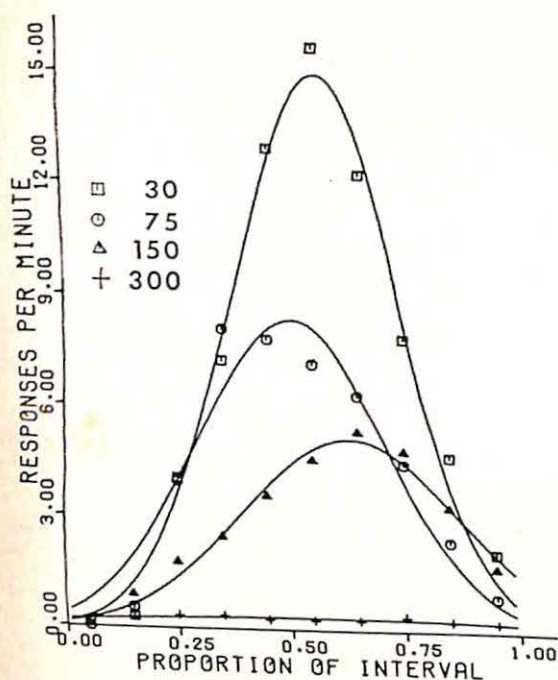


FIGURE 11. Rate of lever pressing for water in Experiment 5c. (The curves were derived from Equation 11.)

largely modulated by the proximity to reinforcement and, when present, discriminative stimuli signaling that proximity (Cole & Litchfield, 1969; Flory, 1969b). In extinction these cues are absent, and both interim and terminal behaviors will blend and exhaust themselves, possibly according to a simple exponential decay function (cf. Azrin et al., 1966; Clark, 1959, 1964).

Experiment 5c

When a rat is exposed to periodic feeding schedules it will consume large quantities of water in the interval between feedings (Falk, 1972). Although similar patterns of excessive drinking have been noted in the rhesus monkey (Schuster & Woods, 1966) and pigeon (Shanab & Peterson, 1969) the intake of a rat is most copious, approaching half of the animal's body weight in a 3-hour session. Explanations in terms of adventitious reinforcement and mediational functions have been ruled out (Falk, 1969). To determine whether the time course of schedule-induced polydipsia corresponds to that of the interim and terminal behaviors investigated in the previous experiments, rats were

exposed to various feedings schedules with water available contingent on a lever press. In a later phase of the experiment, a wooden block was suspended in the cage, and rate of gnawing was measured.

Method. Four naive albino rats, approximately 100 days old, were exposed to feeding schedules of 30-, 75-, 150-, and 300-sec according to a Latin square design. Each depression of a lever delivered .08 ml of water to a reservoir near the food cup; food was not delivered until 5 sec had elapsed without a lever press. After 15 sessions at each of the four feeding intervals, the lever was removed and a small block of pine wood was suspended from a microswitch above the ceiling. Manipulation of the wood sent a train of 10 Hz pulses to the counters for the duration of the activity. With the wood available, the food pellets were delivered every 75 sec, but not within 5 sec of any contact with the wood. After 12 sessions the water was reintroduced, and another 10 sessions were conducted.

Results and discussion. Schedule-induced polydipsia was observed; under all but the longest interfood intervals, the rats consumed several times over their normal daily intake of water. One of the four rats showed regularly increasing response rates up until the moment of reinforcement, despite the DRO contingency employed to minimize superstitious behavior (cf. Segal, 1969). Its data are deleted from further analyses. The behavior of the other three rats is well represented by the averages shown in Figure 11. Overall rate of lever pressing for water was greatest for short intervals and decreased as the interval was lengthened (see

TABLE 4
EXPERIMENTAL CONDITIONS, AVERAGE RESPONSE RATES, AND PARAMETER VALUES FOR EQUATION 11 IN EXPERIMENT 5c

Condition	Response rate	H	S	M
Feeding condition				
30 sec	9.5	14.8	.26	.56
75 sec	6.1	8.3	.29	.50
150 sec	4.3	5.1	.33	.63
300 sec	.3	.3	.75	.51
Drinking (wood present)	3.3	7.1	.27	.56
Gnawing (water absent)	2.2	4.2	.30	.59
Gnawing (water present)	.5	1.1	.23	.50

Note. Response rates are measured as lever presses per minute or seconds gnawing per minute.

Table 4). Lever pressing was maximal midway through the interval, with the variance around .55 showing no systematic relation to the interpellet interval. The curves through the data points of Figure 11 are normal (Equation 11); values of the parameters are listed in Table 4.

When the pine block was introduced, a considerable amount of gnawing took place. Its time course was similar to that of schedule-induced polydipsia. When the water was reintroduced, gnawing decreased substantially, with the maximum rate occurring earlier in the interval. Rate of drinking in the presence of the wood was depressed, with the maximum rate occurring later in the interval.

Schedule-induced polydipsia is often characterized as a "postpellet phenomenon." The present experiments belie that description and show that, while there is some postpellet drinking, the maximum rate of drinking occurs halfway through the interpellet interval. When a block of pine wood is available to the rats, they divert some of the time that would have been occupied drinking water to manipulation and chewing of the wood, with wood gnawing tending to occur slightly earlier in the interval than water drinking.

Both schedule-induced polydipsia and gnawing are reasonably well described by normal curves which account for approximately 94% of the data variance. As rate of reinforcement increased, the maximum of the curves (H) increased, the standard deviation ($S/\sqrt{2}$) decreased, and the midpoint (M) remained unchanged.

This was the first time Equation 11 was invoked to describe data. As the number of steps in a series latency mechanism increases, the resultant probability density function approaches the normal. This follows from the central limit theorem. By the same token, the normal distribution is the limiting form of many different types of mechanisms which might be hypothesized to underlie the obtained distributions. While the obtained data might have modeled with four- and five-step general gamma processes, there would be little point to the exercise, for so

TABLE 5
FIXED-INTERVAL AND PARAMETER VALUES FOR
EQUATION 11 FOR DATA FROM
VARIOUS SOURCES

Source	Fixed interval (sec)	H	S	M
Hawkes and Shimp (1975)	5	180.0	.566	.83
Catania and Reynolds (1968)	30	106.0	.488	1.21
	50	83.3	.488	1.06
	200	46.5	.488	.81
Experiment 2a	60	89.4	.406	.92
Dukich and Lee (1973)	60	55.0	.347	1.05
	120	55.0	.428	1.24

many parameters would dilute conviction in any interpretation of that model.

Why is the normal distribution, rather than a two- or three-step general gamma distribution, appropriate for these data? I suspect the answer lies in the specificity of the measured response. Where it is quite general, as was the case for the experiments reported in the first part of this review, early steps in the gamma process are likely to be captured as measured responses. While subsequent responses may be chained to those early responses and appear normally distributed, it is the earliest behaviors, with the shortest time constants, that contribute most heavily to the final shape of the distribution.

As the set of measured behaviors becomes more specific, we should expect a sharpening of the distribution. If the behaviors occur early in the interval, they should be unalloyed two- or three-step gamma processes. If they occur late in the interval, with many intervening steps in the chain, they should be best described as normal distributions. In the remainder of this article we examine specific responses occurring late in the interval and determine the shape of the distribution which best describes them.

APPLICATION OF THE MODEL TO OPERANT AND RESPONDENT BEHAVIORS

It was noted that one of the four rats in the previous experiment behaved atypically, ingesting water at an increasing rate through the interval. Other investigators have noticed that "schedule-induced" drinking may shift

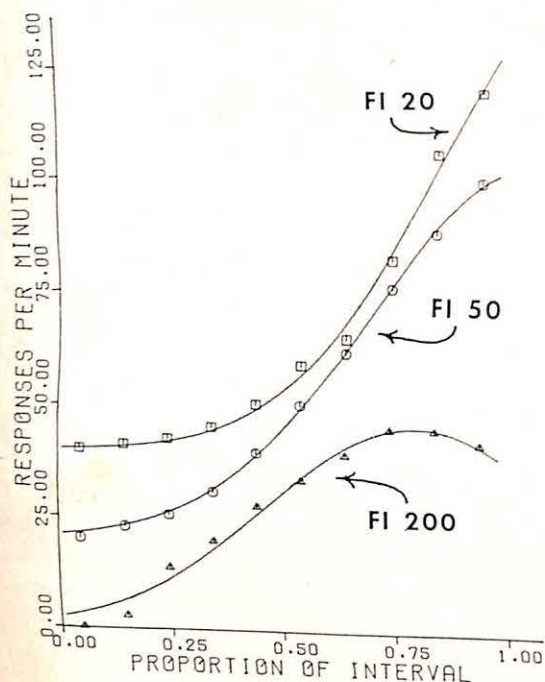


FIGURE 12. Rate of key pecking by pigeons on various fixed-interval schedules. (The top curve, fixed interval 20 sec, is displaced upward by 40 responses per minute, while the middle curve, fixed interval 50 sec, is displaced upward by 20 responses per minute. From Catania & Reynolds, 1968.)

toward the end of the interval and take on the "characteristic form of a fixed-interval scallop." Segal (1972) has suggested that this shift may be characteristic of the genesis of many operant performances—"induction" of some response as an interim behavior, a consequent increase in the probability of contiguity with an ensuing reinforcer, and eventual entrapment as a reinforced, operant response. This shift in the controlling contingencies should be reflected in a shift in the mode and shape of the response distribution. In the previous section we described the distribution of induced behaviors with a normal curve. What is the nature of the distribution of the operant behaviors: What is the mathematical form of a "typical fixed-interval scallop"? In the following paragraphs we select from the literature data on which to base an answer.

In a monograph published in 1968, Catania and Reynolds reported an extensive study of responding by pigeons on variable-interval and fixed-interval schedules. Figure 12 shows the rate of responding as a function

of time through the interval, for fixed intervals of 30, 50, and 200 sec. The curves through the points are normal with parameters given in Table 5. Shown in Figure 13 are data from Hawkes and Shimp (1975) for pigeons responding on an fixed-interval 5-sec schedule. Although these data were collected by different investigators and were generated on a very short fixed-interval schedule, the same normal function which fit Catania and Reynolds's (1968) data is also appropriate for Hawkes and Shimp. The performance of rats on fixed-interval schedules is similar to that of pigeons. Figure 14 shows data collected by Dukich and Lee (1973) studying the responding of rats on fixed intervals of 60 and 120 sec. The parameters are listed in Table 5.

In all cases, behavior is well represented by the left tail of a normal curve, which accounts for 99.5% of the data variance. The present analysis suggests that the form of the typical fixed-interval scallop—the smoothly accelerating curve drawn by a cumulative recorder—is a cumulative normal function. We now turn our attention to the parameters which describe that function.

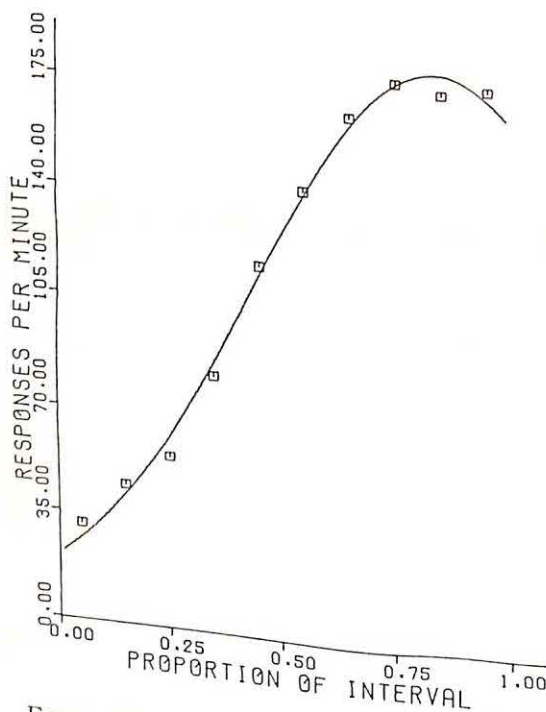


FIGURE 13. Key-pecking rate on an fixed-interval 5-sec schedule. (From Hawkes & Shimp, 1975.)

In Figure 15 I have plotted the product of M and T as a function of T . M may be taken as a psychological judgment of the duration of the interval, and multiplication by T returns it to a nonnormalized form. The data from Experiment 2a and from Catania and Reynolds (1968) fall along a straight line with a slope of .8. The data from Hawkes and Shimp (1975) fall below that line.² On the strength of the former data, we suggest that the psychophysical scale of duration for pigeons is a power function with an exponent of .8. Catania (1970) came to the same conclusion from an analysis of behavior on other reinforcement schedules, as did DeCasper and Zeiler (1974). The fractional slope of the function indicates that pigeons typically overestimate short intervals and underestimate long ones. This tendency is reflected in the truncation of the behavior during a period of accelerating responding on short fixed-interval schedules, and during a period of decelerating responding on longer schedules. For the present

² The differences which exist, such as the larger value of S , the smaller value of M , and the larger intercept at $t = 0$, may be due in part to the use of a 2.5 sec time-out after reinforcement.

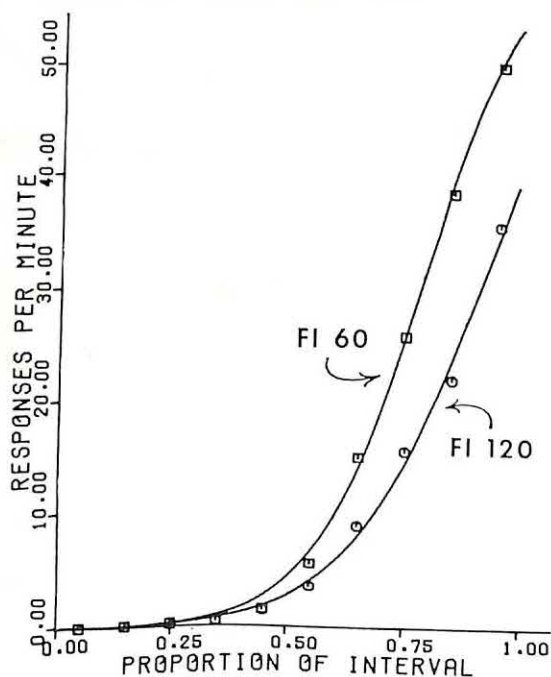


FIGURE 14. Rate of lever pressing by rats on a fixed-interval 60-sec schedule (squares) and a fixed-interval 120-sec schedule (circles). (From Dukich & Lee, 1973.)

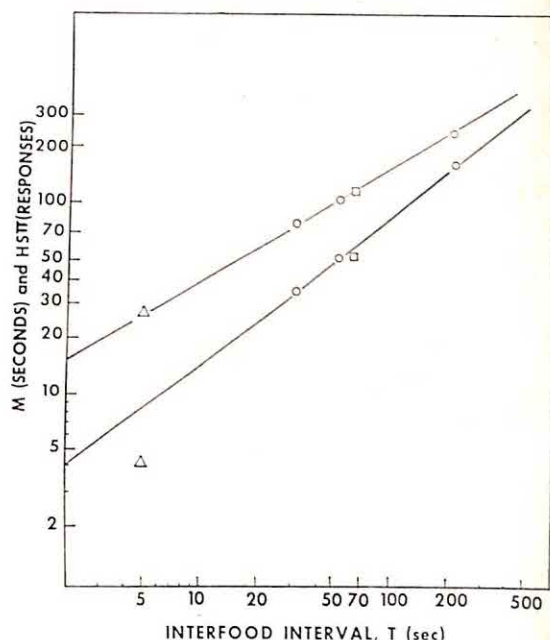


FIGURE 15. Top line: The theoretical maximum number of responses induced by food as a function of the interfood interval. (Data are from Catania & Reynolds, circles; from Experiment 2a, squares; and from Hawkes & Shimp, triangles. The slope of the line is .6.) Bottom line: The relation between interfood interval and midpoint of the normal curve (M). (The data from Catania & Reynolds, circles, and from Experiment 2a, squares, fall along a straight line with a slope of .8, while the data from Hawkes & Shimp, triangles, fall below that line.)

data, estimation is most accurate at a fixed-interval value of about one minute, where reinforcement interrupts pigeons at their maximal rate.

The parameters for the rats are not shown in Figure 15. Their behavior was markedly different from that of pigeons—the maximum rate, H , remained the same for FI 60 and FI 120, while M increased, rather than decreased with T . Since these data are from the control conditions of but one study, they should not be weighed too heavily. If they are replicated, these qualitative differences may implicate an important distinction drawn by Seligman (1970) between “prepared” responses such as pecking, induced by and closely bound to reinforcement, and “unprepared” responses, such as lever pressing, that are more loosely coupled to reinforcement.

The existential status of the right, descending limbs of the truncated normal curves is

uncertain. When reinforcement is only occasionally withheld on fixed-interval schedules, rate does not usually drop off as quickly as would be predicted by the present formulation (cf. Staddon & Innis, 1969; Zimmerman, 1971). A higher ordered exponential equation would provide the necessary "asymmetrical normal curve" shape. Unfortunately the necessary data have not been collected to evaluate such a model. When reinforcement is withheld on a more regular basis (e.g., Catania, 1970; Myerson, 1972), response rate does drop off in a symmetric, normal fashion. Equation 11 therefore seems appropriate for well-learned behaviors such as fixed-interval responding and discriminated extinction. Other models, such as a simple exponential decay function, may apply to nondiscriminated extinction (Clark, 1959, 1964; Furumoto, 1967).

To calculate the maximum number of responses evoked by a reinforcer, one integrates the normal curve from $-\infty$ to $+\infty$ and finds it to be, quite simply, $HS\pi$. In Figure 15 I have plotted $HS\pi$ as a function of T . The equation

$$HS\pi = kT^{.6}, \quad [12]$$

whose locus appears as a solid line, represents well Catania and Reynolds' (1968) data (circles), Experiment 2a data (square), and Hawkes and Shimps' (1975) data (triangle). It should also be mentioned that the unnormalized value of S is approximately proportional to T , and H is a power function of T with an exponent of $-.4$. It is not yet clear which of these relations will turn out to be fundamental, and which derivative.

To predict the average obtained response rate on fixed-interval schedules, one must multiply $HS\pi$ by the area under the curve between $t = 0$ and $t = 1$ and then divide by T . Figure 15 shows us, however, that the normal curve was centered around 1 only for fixed intervals of approximately 60 sec: For shorter schedules, the animals overestimated the time, thereby truncating the normal curve before maximum rate was attained. As the interval grows smaller, less and less of the maximum possible rate is attained. This attenuation might be termed "inhibition by reinforcement" (cf. Gormezano & Moore, 1969; Hovland, 1936). The mechanism of

such "inhibition" is simply the nonlinear psychophysical scale of duration demonstrated in Figure 15.

The present model throws additional light on some regularities in behavior noted by Nevin (1974). Nevin measured response rates on two alternating variable-interval schedules in the presence of explicit competing responses (a fixed-interval schedule on a second key). Over a range of different reinforcement probabilities for the variable interval schedules, he found

$$R_1 = q R_2^s, \quad [13]$$

rate on one schedule was a power function of rate on the other. Rates on the variable-interval schedules were greatest at the beginning of the fixed interval and decreased as the interval progressed and as fixed-interval responding became predominant. Their time course looked much like that of interim behaviors (Figure 1); the analogy has some face validity and will be examined.

Let us simplify Equation 1 by ignoring the inhibition of reinforcement:

$$R_i = A_i e^{-t/C_i}. \quad [14]$$

Response rates on each of the two variable-interval schedules should approximate exponential decay functions of time through the interval. To determine the relation between response rates, we may compare rates only at the same point in time and thus collapse Equation 14 over t :

$$-C_1 \ln(R_1/A_1) = t = -C_2 \ln(R_2/A_2). \quad [15]$$

Rearrangement yields

$$R_1 = A_1 \left(\frac{R_2}{A_2} \right)^{C_2/C_1}, \quad [16]$$

a form equivalent to Equation 13, with

$$s = C_2/C_1, \quad [17]$$

and

$$q = A_1/A_2^s. \quad [18]$$

While Nevin's (1974) Equation 13 thus follows from the present theory, independent quantitative predictions of his data are difficult, because the present analysis provides no basis for predicting the nature of changes

in C with changes in the probability of reinforcement for interim behaviors (i.e., variable-interval schedule responding). If we assume, however, that A is proportional to the rate of reinforcement, we may use Nevin's values of s in Equation 18 to predict his obtained values of q . The unknown constant of proportionality enters Equation 18 in a nonlinear fashion, causing the prediction for individual animals to be in error by the factor $k^{(1-s)}$. Nonetheless, approximate prediction is possible, and the scattergram of obtained versus predicted intercepts (Figure 16) shows it to be effective.

The relevance of the present analysis to Nevin's data is important in light of the assumptions that were made in the course of the analysis. Key pecking on multiple variable-interval schedules, maintained by contingent reinforcement, was successfully treated as an "interim" behavior. Such responding shares with interim behaviors the same relatively homogenous probability of reinforcement over time, and, like interim behaviors, suffers competition from periodically occurring reinforcers. The use of Equations 1 and 10 should not, therefore, be thought to be restricted to schedule-induced, or interim behaviors but may be taken as a model of the changes in any tonic, or background, behavior in the context of periodic occurrences of reinforcers which elicit competing behavior.

Application to Respondent Behavior

While the paradigms which generated much of the data in the present experiments would be classified as respondent, or classical conditioning, it is useful to look at a more traditional instance of respondent behavior from the perspective of the present theory. Pavlov (1960; pp. 95, 96) reported several experiments involving "inhibition of delay"; conditioned stimuli of 3-min duration were presented at intervals of 12 min. The unconditioned stimulus was the injection of a weak acid solution into a dog's mouth. Although some of the trials involved compound stimuli, Pavlov reported the number of drops of saliva emitted every 30 sec during 20 control trials. I have averaged these data, transformed them into a rate measure, and graphed them in Figure 17. A normal curve

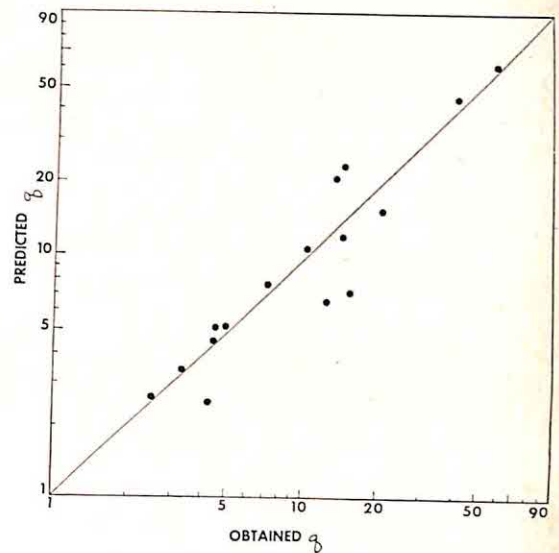


FIGURE 16. Obtained versus predicted values of q , the intercepts of Nevin's (1974) power functions. (The predictions followed from Equation 18, with A assumed proportional, $k = 150$, to the reinforcements per minute. The predictions are for individual subjects from each condition, with the exception of Pigeon 60 Condition 4, for whom no slope could be determined.)

with a mean of .95 and a standard deviation of .43 accounts for practically all of the data variance.

Equations 10 and 11 are also reasonable descriptors of the behavior occurring during much shorter conditioned stimulus-unconditioned stimulus intervals. Smith (1968) classically conditioned the nictitating-membrane response of rabbits, employing intervals between 125 and 1,000 msec. Third-order exponential functions described the data from the three shortest intervals and a normal curve described the data from the longest interval. Approximately 95% of the data variance was accounted for by those models.³

GENERAL DISCUSSION

The study of the general activity of pigeons during the interval between feedings generated regular data conducive to representation

³ An additional transformation was necessary to permit these models to fit Smith's data: 90 msec were arbitrarily subtracted from the abscissa before Equations 10 and 11 were fit. This latency should be treated as an additional free parameter and might be interpreted as an estimate of the absolute refractory period.

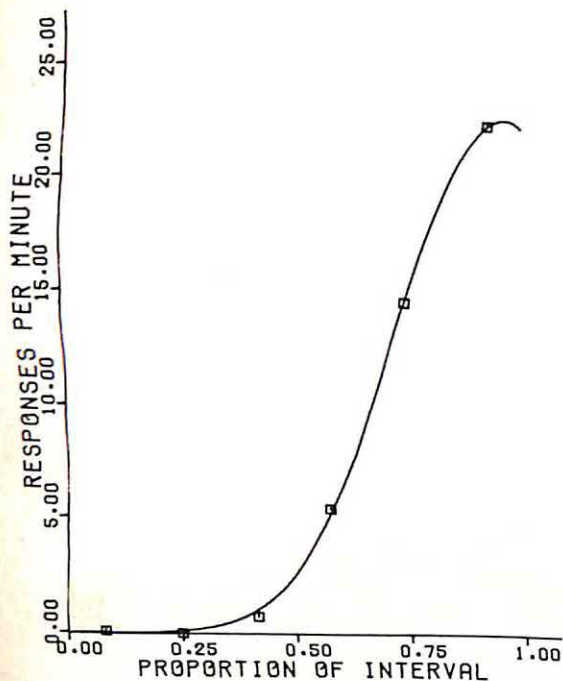


FIGURE 17. Rate of salivation (drops per minute) as a function of time within a 3 min conditioned stimulus (data from Pavlov, 1960).

by a simple mathematical model (Equation 1). The two exponential decay functions of the model were identified as a short-term inhibitory process reflecting the remoteness of one food delivery from the next and a longer term process reflecting the rise of competing ("terminal") behaviors.

Whereas this simple model provided an adequate account of general activity, the addition of a third exponential was necessary to accommodate the interim behaviors reported by Staddon and Simmelhag (1971), and it improved the fit to some other data. This third, very fast exponential permitted the early portions of the curves to show positive acceleration and was attributed to a carry-over of consummatory responses or orienting responses into the beginning of short intervals. Although responses occurring toward the middle and end of an interval could be represented by three or four interacting exponentials, the normal function required fewer parameters and provided a superior fit to the data. The normal curve is the limiting distribution when many exponential decay functions interact.

Changes in the parameters of the model showed a gratifying concordance with changes

in independent variables. The theoretical maximum, A , was proportional to the feeding rate, while H was a power function of the feeding rate. The time constants of the component exponential processes showed regular changes with changes in rate of feeding and a consistent (inverse) relationship with each other. The parameters of the normal curve were either relatively invariant or changed smoothly with changes in the interfood interval.

Equations 1 and 10 may be interpreted as the impulse response curves of second and third order control systems. In this construction, response rate is presumed proportional to the deviation of the system from homeostasis. This deviation is brought about by an impulse—in the present case, food delivery. It may be assumed either that the measured response drives the system back to equilibrium (a traditional servo-system with proportional negative feedback) or simply that the return occurs as an exponential function of time. The latter assumption is more congenial to the behavior in question. It is also necessary to assume either inertia in the system or an exponential delay involved in communicating the impulse to the response mechanism. The latter interpretation is adopted. (Still other control devices are modeled by second order systems, e.g., devices with proportional and integral control, but these seem further removed from the data, cf. Baylis, 1966).

The mapping of induced behaviors onto control systems, via Equations 1 and 10, brings into focus some behavioral regularities and provides indices for other behavioral changes. The damping ratio ($\xi = \sqrt{C/I/2}$) governs the speed with which the system returns to equilibrium. This index changed in the appropriate manner in Experiment 2 and across Experiments 1a and 1b. The constancy of the normalized resonant frequency (ω) provided a picture of organisms finely tuned to the temporal regularities in their environment. At the same time, however, the dependence of this basic system parameter on the input frequency ruled out both a simple reflexive interpretation of the data and the general usefulness of simple linear control systems, which predicate a reflexive system having input-invariant param-

eters! Further experiments should be conducted before a cybernetic interpretation is discarded. But one of the most severe impediments to this construction is its interpretation of a low response rate before food delivery as a return to equilibrium. If that feeding were omitted, response rate would soon increase to record highs, a phenomenon beyond the purview of simple control theory. Although such a rebound might be accommodated by control systems with predictive elements, the complexity of those models is not justifiable with the present limited data base.

An unramified stochastic process, such as was presented earlier, fares little better than control theory as an interpretation of the fundamental equations. The constancy of the number of inferred parallel processes in Experiments 1a and 1b was encouraging. It is possible that simple modification of the organism's environment, such as a change in chamber size, will have direct effects on the index. It is possible that the steps in the chain will be correlated with the perception of and reaction to environmental stimuli and thus tie the present model into stimulus-sampling theories of behavior acquisition (Estes, 1959; Kintsch, 1970).

But there are problems with the simple stochastic interpretation. The number of hypothetical processes was constant in no other parts of these experiments, and the changes that occurred under drugs were unenlightening. Especially damaging is the constancy of the product of C and I over large changes in feeding rate and insult by various drugs. This powerful invariance in the data would go unnoticed for there is no index in series latency mechanics which involves the product of C and I . Finally, the construal of individual responses from homogenous sets of behavior, such as pacing or drinking, as independent outcomes of parallel processes seems unsympathetic to the data.

General activity comprises many behaviors which hang together as unitary phenomena. The stochastic model may be modified to preserve this molar structure of the data. Assume the gamma distribution represents a "behavioral state" of the organism during which time a particular activity, such as

polydipsic drinking, occurs at a uniform rate. It is called a "behavioral" state to avoid connotations of a "drive" state, such as "thirst." Let us assume the presence of another state, represented by a gamma distribution with the same values of A and I but a different value of C (C'). Correlated with this state are other behaviors, such as running in a wheel, which are suppressed ("displaced") by the behaviors of the first state. The probability of being in the second state at a given point in time (and therefore the average rate at that time) will be equal to the difference of the two equations:

$$R = A(e^{-t/C} - e^{-t/C'}). \quad [19]$$

This behavioral state development seems most in accord with observed behaviors. Furthermore, it explains quantitative properties of data collected by Staddon and Ayres (in press) and analysed by them and by Staddon (in press) in a state theoretic manner. Their analysis, which motivated the above development, showed that the duration of one activity (e.g., running) was negatively correlated with the time of offset of a previous activity (e.g., drinking). They took the significant negative correlations, along with data from experiments in which one of the activities was prevented, as evidence *against* a simple chaining interpretation (i.e., drinking leads to running) and *for* central timing of each of the states. Inspection of Equation 19 shows that the total amount of time in a state will indeed be negatively correlated with variations in the time constant governing entry into that state (C'). The machinery of generalized gamma processes may thus be retained, but the emphasis upon the termination of a state is now shifted to the behavior occurring during a state. The "latent period" is revealed by the measured responses, which are taken to signify the outcome of competition among parallel behavioral states.

There are implications to the present research that go beyond the provision of an adequate model for the time course of behavior and the leverage on related problems which that provides. Four historically distinct phenomena—general activity, schedule-induced behaviors, operants, and respondents

have been encompassed by a single mathematical formulation. While there are qualitative distinctions among these various behaviors, they have a common foundation in Equations 1, 10, and 11. If future research shows that multiple behaviors within an interval share some parameters (e.g., A, corrected for response units) and exert predictable effects on the parameters of other behaviors (as the introduction of an operant in Experiment 2a increased the value of C), the theoretical usefulness of those distinctions will fade.

What then will remain as important discriminanda among behaviors? Operants are said to be emitted while respondents are elicited. But "emission" cannot be taken to imply lack of prior causal stimuli (Staddon, 1973); all operants must be "elicited" by stimuli of some sort. If an external discriminative stimulus is not provided, time will serve. And time serves operants and respondents similarly, for we have seen that the same mathematical function of time describes both types of behavior.

Respondents are said to extinguish more quickly than operants and to suffer greater decrement under conditions of partial reinforcement. But that distinction is confounded by the historic difference between the response systems studied within the operant and respondent paradigms. When similar systems are studied, similar partial-reinforcement effects may be found (e.g., Williams & Williams, 1969). In the experiments modeled in this article, many key pecks and many drops of saliva occurred during the intervals between feedings. Which was on a thinner schedule of partial reinforcement is moot. The same equation describes the time course of each and provides theoretical asymptotes for each. Only when the scale factors for various behaviors receive relative calibration (e.g., by the measurement of concomitant changes under manipulations of drugs, deprivation, or amount of reinforcement) can meaningful comparisons be made. And only then can the apparent differences in resistance to extinction in discrete-trial experiments be fully understood.

Some behaviors are more malleable than others; typical respondents are not so readily

shaped with differential reinforcement and might therefore justify distinction from operants. But the Brelands (1966) have alerted us to the limitations on our ability to shape operant behaviors. Moore (1973) suggests that the topography of operants is closely related to forms of behavior induced by "reinforcement"; contingencies merely change the location of induced behaviors. Hearst and Jenkins (1975) provide a possible mechanism for that spatial control in their analysis of "sign-tracking" (cf. Neuringer & Neuringer, 1974; Schneirla, 1959). Certainly more research is necessary, but the presently available data suggest that sign-tracking, schedule-induced arousal, and the present model of temporal control may provide the basic "laws" for a theory of behavior more powerful than any founded on the law of effect.

The asymptotic level of general activity, schedule-induced behaviors, and operants increases with increases in the frequency of reinforcement. This systematic change is taken as evidence for an intervening state of "arousal." The construct has much in common with the "anticipation-invigoration mechanism" of Cofer and Appley (1967), the "emotion" of Segal (1972), the "central-motive state" of Bindra (1972), and the "arousal" of Wayner (1970), Valenstein, Cox, and Kakolewski (1970), Glickman (1973), and Fentress (1973). It is taken here to be a general and nonspecific activator of behaviors, functioning essentially as the gain-control of an amplifier. An essential problem for this and all similar theories is that of specifying which behaviors will be amplified—what functions as the channel selector of the organism?

While temporal cues govern the transition between behavioral states, situational cues determine the nature of the state selected. Within a range dictated by its ecological niche, the history and immediate environment of the organism become important factors in determining the types of behavior that will be induced. The plasticity of the induced behaviors and the amount and kind of stimulus support necessary to release them from an aroused organism remain important and largely unexplored topics for research.

An important question remains: What in

the present theory distinguishes "normal," "operant," or "terminal" behaviors from "interim," "adjunctive," or "displacement" behaviors? Two aspects of the incentive-motivator will account for the distinction: (a) The incentive is available only at the end of the interval; behavior appropriate to it will therefore be elicited toward the end of the interval, and only then; and (b) Adequate contact with the incentive will satiate the organism, thereby turning off the arousal.⁴

Elements of the present theory have been suggested before (cf. Bindra, 1972, Staddon, 1972, on temporal versus situational control; Moore, 1973, on Pavlovian mechanisms in operant behavior; Fentress, 1973, on the nonspecificity of the central motive state of arousal). They may seem too simple and incomplete to account for the data under purview; but consider: The theory suggests that a state of elevated arousal is generated by events traditionally referred to as reinforcers. The arousal increases the vigor of all responses elicited by situational stimuli, both interoceptive and exteroceptive, acting in concert with temporal stimuli. The elicitation of behavior is quite general, with those behaviors favored by the organism's phylogenetic or ontogenetic history and present drive condition being predominant. Since many of these behaviors do not abate the organism's arousal, they may continue at "excessive" levels. But as time wears on, behaviors appropriate to the appetited stimulus will increase in frequency according to equations such as 10 and 11. All of the "fine control" of behavior attributed to law of effect strengthening may be reducible to the time-contingent elicitation of approach to stimuli that are correlated with reinforcement.

The present theory repudiates the traditional distinctions between classes of behavior and questions the future usefulness of the law of effect. That law was historically necessary, for it established "reinforcement" as a firm temporal reference point from which to view the immediately preceding "terminal" behaviors. With the present

model of temporal control, we may trace the course of all behaviors that are referenced to signals of forthcoming incentives, such as feedings. "Reinforcement" and its retroflex action may be dispatched.

REFERENCES

- Amsel, A. Frustrative nonreward in partial reinforcement and discrimination learning: Some recent history and a theoretical extension. *Psychological Review*, 1962, 69, 306-328.
- Amsel, A. Behavioral habituation, counterconditioning, and a general theory of persistence. In A. H. Black & W. F. Prokasy (Eds.), *Classical conditioning II: Current theory and research*. New York: Appleton-Century-Crofts, 1972.
- Amsel, A., & Work, M. S. The role of learned factors in "spontaneous" activity. *Journal of Comparative and Physiological Psychology*, 1961, 55, 527-532.
- Azrin, N. H., Hutchinson, R. R., & Hake, D. F. Extinction-induced aggression. *Journal of the Experimental Analysis of Behavior*, 1966, 9, 191-204.
- Baylis, L. E. *Living control systems*. San Francisco: Freeman, 1966.
- Bindra, D. Neuropsychological interpretation of general activity and instrumental behavior. *Psychological Review*, 1968, 75, 1-22.
- Bindra, D. The interrelated mechanisms of reinforcement and motivation, and the nature of their influence on response. In W. J. Arnold & D. Levine (Eds.), *Nebraska Symposium on Motivation* (Vol. 18). Lincoln: University of Nebraska Press, 1970.
- Bindra, D. A unified account of classical conditioning and operant training. In A. H. Black & W. F. Prokasy (Eds.), *Classical conditioning II: Current research and theory*. New York: Appleton-Century-Crofts, 1972.
- Bindra, D., & Palfai, T. Nature of positive and negative incentive-motivational effects in general activity. *Journal of Comparative and Physiological Psychology*, 1967, 63, 288-297.
- Blough, D. S. New test for tranquilizers. *Science*, 1958, 127, 586-587.
- Bolles, R. C. *Theory of motivation*. New York: Harper & Row, 1967.
- Bolles, R. C. Interactions with motivation. In M. Marx (Ed.), *Learning: Interactions*. London: Macmillan, 1970. (a)
- Bolles, R. C. Species-specific defense reactions and avoidance learning. *Psychological Review*, 1970, 77, 32-48. (b)
- Breland, K., & Breland, M. The misbehavior of organisms. *American Psychologist*, 1961, 16, 681-684.
- Breland, K., & Breland, M. *Animal behavior*. New York: Macmillan, 1966.
- Brown, P. L., & Jenkins, H. M. Auto-shaping of the pigeon's key-peck. *Journal of the Experimental Analysis of Behavior*, 1968, 11, 1-8.

⁴ This feature is missing when the reinforcer is intracranial stimulation of the brain.

- Campbell, B. A., & Sheffield, F. D. Relation of random activity to food deprivation. *Journal of Comparative and Physiological Psychology*, 1953, 46, 320-322.
- Capaldi, E. J. Partial reinforcement: A hypothesis of sequential effects. *Psychological Review*, 1966, 73, 459-477.
- Capaldi, E. J. An analysis of the role of reward and reward magnitude in instrumental learning. In J. H. Reynierse (Ed.), *Current issues in animal learning*. Lincoln: University of Nebraska Press, 1970.
- Carlton, P. L. Brain-acetylcholine and inhibition. In J. T. Tapp (Ed.), *Reinforcement and behavior*. New York: Academic Press, 1969.
- Catania, A. C. Reinforcement schedules and psychophysical judgments. In W. N. Schoenfeld (Ed.), *The theory of reinforcement schedules*. New York: Appleton-Century-Crofts, 1970.
- Catania, A. C., & Reynolds, G. S. A quantitative analysis of the responding maintained by interval schedules of reinforcement. *Journal of the Experimental Analysis of Behavior*, 1968, 11, 327-383.
- Clark, F. C. Some quantitative properties of operant extinction data. *Psychological Reports*, 1959, 5, 131-139.
- Clark, F. C. Effects of repeated VI reinforcement and extinction upon operant behavior. *Psychological Reports*, 1964, 15, 943-955.
- Cofer, C. N., & Appley, M. H. *Motivation: Theory and research*. New York: Wiley, 1967.
- Cohen, P. S., & Looney, T. A. Schedule-induced mirror responding in the pigeon. *Journal of the Experimental Analysis of Behavior*, 1973, 20, 395-408.
- Cole, J. M., & Litchfield, P. M. Stimulus control of schedule-induced aggression in the pigeon. *Psychonomic Science*, 1969, 17, 152-153.
- DeCasper, A. J., & Zeiler, M. D. Time limits for completing fixed ratios. III. Stimulus variables. *Journal of the Experimental Analysis of Behavior*, 1974, 22, 285-300.
- Dukich, A. D., & Lee, A. E. A comparison of measures of responding under fixed-interval schedules. *Journal of the Experimental Analysis of Behavior*, 1973, 20, 281-290.
- Estes, W. K. The statistical approach to learning theory. In S. Koch (Ed.), *Psychology: A study of a science* (Vol. 2). New York: McGraw-Hill, 1959.
- Falk, J. L. Conditions producing psychogenic polydipsia in animals. *Annals of the New York Academy of Sciences*, 1969, 157, 569-593.
- Falk, J. L. The nature and determinants of adjunctive behavior. In R. M. Gilbert & J. D. Keehn (Eds.), *Schedule effects: Drugs, drinking, and aggression*. Toronto, Canada: University of Toronto Press, 1972.
- Fentress, J. C. Specific and nonspecific factors in the causation of behavior. In P. P. G. Bateson & P. H. Klopfer (Eds.), *Perspectives in ethology*. London: Plenum, 1973.
- Flory, R. K. Attack behavior as a function of minimum inter-food interval. *Journal of the Experimental Analysis of Behavior*, 1969, 12, 825-828. (a)
- Flory, R. K. Attack behavior in a multiple fixed-ratio schedule of reinforcement. *Psychonomic Science*, 1969, 16, 156-157. (b)
- Furumoto, L. *Studies of experimental extinction in the pigeon*. Unpublished doctoral dissertation, Harvard University, 1967.
- Glickman, S. E. Responses and reinforcement. In R. Hindle & J. Stevenson-Hindle (Eds.), *Constraints on learning: Limitations and predispositions*. New York: Academic Press, 1973.
- Gormezano, I., & Moore, J. W. Classical conditioning. In M. Marx (Ed.), *Learning: Processes*. London: Macmillan, 1969.
- Hawkes, L., & Shimp, C. P. Reinforcement of behavioral patterns: Shaping a scallop. *Journal of the Experimental Analysis of Behavior*, 1975, 23, 3-16.
- Hearst, E., & Jenkins, H. M. *Sign-tracking: The stimulus-reinforcer relation and directed action*. Austin, Texas: Psychonomic Society, 1975.
- Hovland, C. I. "Inhibition of reinforcement" and phenomena of experimental extinction. *Proceedings of the National Academy of Science, Washington*, 1936, 22, 430-433.
- Jenkins, H. M. Sequential organization in schedules of reinforcement. In W. N. Schoenfeld (Ed.), *The theory of reinforcement schedules*. New York: Appleton-Century-Crofts, 1970.
- Jensen, C., & Fallon, D. Behavioral aftereffects of reinforcement and its omission as a function of reinforcement magnitude. *Journal of the Experimental Analysis of Behavior*, 1973, 19, 459-468.
- Keehn, J. D. Schedule-dependence, schedule-induction, and the law of effect. In R. M. Gilbert & J. D. Keehn (Eds.), *Schedule effects: Drugs, drinking, and aggression*. Toronto, Canada: University of Toronto Press, 1972.
- Kintsch, W. *Learning, memory, and conceptual processes*. New York: Wiley, 1970.
- Konorski, J. *Integrative activity of the brain: An interdisciplinary approach*. Chicago: University of Chicago Press, 1967.
- Levitsky, D., & Collier, G. Schedule-induced wheel running. *Physiology and Behavior*, 1968, 3, 571-573.
- Margules, D. L., Margules, A. S. The development of operant responses by noradrenergic activation and cholinergic suppression of movements. In J. D. Maser (Ed.), *Efferent organization and the integration of behavior*. New York: Academic Press, 1973.
- McFarland, D. J. *Feedback mechanisms in animal behavior*. New York: Academic Press, 1971.
- McGill, W. J. Stochastic latency mechanisms. In R. D. Luce, R. R. Bush, & E. Galanter (Eds.), *Handbook of mathematical psychology* (Vol. 1). New York: Wiley, 1963.
- McGill, W. J., & Gibbon, J. The general-gamma distribution and reaction times. *Journal of Mathematical Psychology*, 1965, 2, 1-18.

- Milsum, J. H. *Biological control systems analysis*. New York: McGraw-Hill, 1966.
- Moore, B. R. The role of directed Pavlovian reactions in simple instrumental learning in the pigeon. In R. A. Hinde & J. Stevenson-Hinde (Eds.), *Constraints on learning: Limitations and predispositions*. New York: Academic Press, 1973.
- Myerson, J. *Trace stimulus control in second-order schedules*. Unpublished doctoral dissertation, Arizona State University, 1972.
- Neuringer, A., & Neuringer, M. Learning by following a food source. *Science*, 1974, 184, 1005-1008.
- Nevin, J. A. Response strength in multiple schedules. *Journal of the Experimental Analysis of Behavior*, 1974, 21, 389-408.
- Pavlov, I. P. *Conditioned reflexes*. G. V. Anrep, (trans.). New York: Dover, 1960.
- Schneirla, T. C. An evolutionary and developmental theory of biphasic processes underlying approach and withdrawal. In *Current theory and research on motivation*. Lincoln: University of Nebraska Press, 1959.
- Schuster, C. R., & Woods, J. H. Schedule-induced polydipsia in the rhesus monkey. *Psychological Reports*, 1966, 19, 823-828.
- Scull, J. W. The Amsel frustration effect: Interpretations and research. *Psychological Bulletin*, 1973, 79, 352-361.
- Segal, E. F. Transformation of polydipsia drinking into operant drinking: A paradigm? *Psychonomic Science*, 1969, 16, 133-135.
- Segal, E. F. Induction and the provenance of operants. In R. M. Gilbert & J. R. Millenson (Eds.), *Reinforcement: Behavioral analyses*. New York: Academic Press, 1972.
- Seligman, M. E. P. On the generality of the laws of learning. *Psychological Review*, 1970, 77, 406-418.
- Shanab, M. E., & Peterson, J. L. Polydipsia in the pigeon. *Psychonomic Science*, 1969, 15, 51-52.
- Sheffield, F. D. A drive-induction theory of reinforcement. In R. N. Haber (Ed.), *Current research in motivation*. New York: Holt, 1966.
- Sheffield, F. D., & Campbell, B. A. The role of experience in the "spontaneous" activity of hungry rats. *Journal of Comparative and Physiological Psychology*, 1954, 47, 97-100.
- Siegel, S., & Domjan, M. Backward conditioning as an inhibitory procedure. *Learning and Motivation*, 1971, 2, 1-11.
- Skinner, B. F. Superstition in the pigeon. *Journal of Experimental Psychology*, 1948, 38, 168-172.
- Slivka, R. M., & Bitterman, M. E. Classical appetitive conditioning in the pigeon: Partial reinforcement. *Psychonomic Science*, 1966, 4, 181-182.
- Smith, M. C. CS-US interval and US intensity in classical conditioning of the rabbit's nictitating membrane response. *Journal of Comparative and Physiological Psychology*, 1968, 66, 679-687.
- Staddon, J. E. R. Temporal effects of reinforcement: A negative "frustration" effect. *Learning and Motivation*, 1970, 41, 227-247.
- Staddon, J. E. R. Temporal control and the theory of reinforcement schedules. In R. M. Gilbert & J. R. Millenson (Eds.), *Reinforcement: Behavioral analyses*. New York: Academic Press, 1972.
- Staddon, J. E. R. On the notion of cause, with applications to behaviorism. *Behaviorism*, 1973, 1, 25-63.
- Staddon, J. E. R. Temporal control, attention, and memory. *Psychological Review*, 1974, 81, 375-391.
- Staddon, J. E. R. Schedule-induced behavior. In W. K. Honig & J. E. R. Staddon (Eds.), *Operant behavior* (Vol. 2). New York: Appleton-Century-Crofts, in press.
- Staddon, J. E. R., & Ayres, S. L. Sequential and temporal properties of behaviors induced by a schedule of periodic food delivery. *Behavior*, in press.
- Staddon, J. E. R., & Innis, N. K. Reinforcement omission on fixed-interval schedules. *Journal of the Experimental Analysis of Behavior*, 1969, 12, 689-700.
- Staddon, J. E. R., & Simmelhag, V. L. The "superstition" experiment: A re-examination of its implications for the principles of adaptive behavior. *Psychological Review*, 1971, 78, 3-43.
- Valenstein, E. S., Cox, V. C., & Kakolewski, J. W. Re-examination of the role of the hypothalamus in motivation. *Psychological Review*, 1970, 77, 16-31.
- Wayner, M. J. Motor control functions of the lateral hypothalamus and adjunctive behavior. *Physiology and Behavior*, 1970, 5, 1319-1325.
- Williams, D. R., & Williams, H. Auto-maintenance in the pigeon: Sustained pecking despite contingent non-reinforcement. *Journal of the Experimental Analysis of Behavior*, 1969, 12, 511-520.
- Zamble, E. Classical conditioning of excitement anticipatory to food reward. *Journal of Comparative and Physiological Psychology*, 1967, 63, 526-529.
- Zamble, E., & Kirkwood, R. Discriminative classical conditioning of excitement anticipatory to food reward. *Psychonomic Science*, 1969, 16, 252-253.
- Zener, K. The significance of behavior accompanying conditioned salivary secretion for theories of the conditioned response. *American Journal of Psychology*, 1937, 50, 384-403.
- Zimmerman, D. W. Rate changes after unscheduled omission and presentation of reinforcement. *Journal of the Experimental Analysis of Behavior*, 1971, 15, 261-270.

(Received May 29, 1974; revision received October 17, 1974)

Arousal, Activation, and Effort in the Control of Attention

Karl H. Pribram
Stanford University

Diane McGuinness
University College, London, and the Hatfield
Polytechnic, Hatfield, England

This review attempts to organize a range of neuropsychological and psychophysiological data on attention. Three separate, but interacting, neural systems are distinguished: One controls *arousal*, which is defined in terms of phasic physiological responses to input. The arousal control circuits center on the amygdala. A second system controls *activation*, which is defined in terms of tonic physiological readiness to respond. The readiness circuits center on the basal ganglia of the forebrain. A third system is discerned which coordinates arousal and activation. This coordinating activity is defined as demanding *effort*. Its circuitry centers on the hippocampus. When arousal, activation, and effort are involved in problem solving, at least two further distinctions can be made. During *categorizing*, arousal precedes activation; during *reasoning*, activation precedes arousal. Finally, the review addresses the question of whether effort in problem solving is to be attributed solely to peripheral muscular factors or whether, in fact, direct monitoring of changes in brain organization (changes in set, attitude, etc.) can be productive of measurable indicators of effort.

In the mid 1960s, R. J. Douglas and K. H. Pribram presented a series of experiments detailing the effects of hippocampal lesions on problem-solving behavior. The results of *these experiments were accounted for* by recourse to the concept of experiencing a change in awareness which was interpreted *as indicating the involvement of "attention."* A continuing interest in problems of attention has been reflected not only in the work reported from our laboratory (a dozen and a half papers) but also in that of many others (see reviews by Mostofsky, 1970; Horn and Hinde, 1970; and Kahneman, 1973). Research has largely been directed to two general paradigms defined by the following *operations: (a) recording the occurrence of physiological or behavioral responses* against a background of monotonous repetition of

sensory events (orienting, vigilance, and habituation) and (b) the invariant (categorizing), or the variable but computable (reasoning), pairing of the outcome of response (reinforcement) to sensory events. An assessment of the data derived from these paradigms (in approximately 200 experiments) has led us to identify three basic attentional control processes: One regulates *arousal* resulting from input; a second controls the preparatory *activation* of response mechanisms; and a third operates to coordinate arousal and activation, an operation that demands *effort*. The first two sections of this review will delineate these attentional control systems; the last section will deal primarily with higher order cortical attentional controls that operate during categorizing and reasoning.

We wish to express our gratitude to Muriel Pagshaw for additional data analysis and for allowing us to present her unpublished work in this manuscript, and to Linda Heider and Barbara Honegger for their patient help. The research involved was supported by National Institute of Mental Health Grant MH-12970-08 and National Institute of Mental Health Career Award MH-15214-13 to the first author.

Requests for reprints should be sent to Karl H. Pribram, Department of Psychology, Stanford University, Stanford, California 94305.

REFLEX ATTENTION, REPRESENTATION, AND THE CONTROL OF AROUSAL

Arousal is said to occur when an input change produces a measurable incrementing of a physiological (e.g., single unit recording of neural potentials; galvanic skin response) or behavioral (e.g., response amplitude of a spinal reflex; frequency of a locomotor response) indicator over a baseline. The types of input change that produce

arousal have been studied extensively and have been labeled by Berlyne (1969) as collative variables. These include sudden changes in intensity to which the organism is unaccustomed, changes in the timing of inputs, and changes in the ground in which a stimulus figure appears, in particular when input is scarce, surprising, complex, and novel. Such collative characteristics also define the concept "information" as it is used in the study of communication systems (see e.g., Brillouin, 1962); thus it is possible to treat organisms subject to arousal as "information processing systems," where input is matched against some residual in the organism of his past experience, or some competence (Miller, Galanter, & Pribram, 1960; Pribram, 1971). Without such matching, there could be no novelty or information nor even a measure of change in intensity.

This knowledge has led to behavioral research which has been addressed to such issues as adaptation level (Helson, 1964), expectancy (Bruner, 1957) and the development of "neuronal models" (Sokolov 1960, 1963), while neuroscientists were able to demonstrate the occurrence of a permanent or semipermanent modifiability of neural tissue (see Pribram and Broadbent, 1970; Horn and Hinde, 1970).

Neural Modifiability

Sharpless and Jasper's (1956) and Sokolov's (1963) classical studies paved the way for an analysis of the requisites for producing behavioral and physiological arousal responses. Especially significant was the implication of information processing systems by Sokolov's findings (1963) that *any* change in a repetitively presented stimulus configuration would elicit responses even when that change was a shortening of the sensory input or the lowering of its intensity. Sokolov concluded that the organism's sensitivity was not just a measure of its threshold but rather that sensitivity was a function of a match between input and some competence, a neuronal "model," or a patterned memory trace developed in the brain as a representation of the experienced stimulus configuration. Any mismatch produces orienting.

Many investigators concerned with modifi-

ability of neural tissue as a consequence of experience seized on the habituation of an initial neural or behavioral arousal reaction as an extremely simple test paradigm that suited this purpose. Although these efforts and results have been reviewed by Horn and Hinde (1970), we will highlight certain issues that are relevant to this presentation.

An approach initiated by Spencer and Thompson (see Groves and Thompson, 1970) involved recording from the spinal cord and illustrated that neither sensory afferents nor motor efferents accounted for habituation of spinal reflex behavior. Further, Thompson discerned two primary systems, two classes of neurons; those that responded more slowly and incremented for relatively long periods before decrementing, and those that rapidly decremented their activity with repetition of the stimulus. The incrementing "monitoring," "sameness," or "arousal" neurons lie more ventrally and medially in the region of the origin of the visceromotor outflow, while the decrementing neurons lie, for the most part, in the dorsal horn of the spinal cord. Convergence of these two systems of neurons was assumed to account for the activity of a third population of ("novelty") neurons whose response characteristics were identical to those demonstrated by the reflex: arousal reaction, habituation, and dishabituation.

Thompson's parallel processing model of arousal, habituation and dishabituation, shared by many other investigators (e.g., Horn, 1970; Segundo & Bell, 1970; Spencer & April, 1970) is simpler than Sokolov's as it depends only on spatial summation of arousal and decrementing. No "neuronal model" of the input is developed, only a decrementing to repetition and an "arousal" proportional to the strength of the input. The model does not account for Sokolov's findings that in the intact organism dishabituation occurs when stimulus duration is *shortened* or stimulus intensity is *diminished*. These phenomena imply a hierarchical process involving storage and subsequent matching of input to store. Horn (Horn & Hinde, 1970) has developed a scheme whereby such a hierarchical process can be realized without any unnecessary assumptions that go beyond

the data obtained by practically everyone; namely that the decrementing observed is due to self-generated synaptic depression and not to the development of inhibition (inhibitory postsynaptic potentials characterized in intracellular recordings by hyperpolarization).

This does not mean, of course, that inhibition fails to occur in many systems. The sensory and motor pathways are replete with circuits that lead to both lateral (i.e., neighborhood) and self-inhibition. As we shall see shortly, habituation becomes organized into patterns in these pathways whether by the mechanisms proposed by Horn or by the action of inhibitory circuits, or both.

Core Brain Arousal System

The rostral extension into the mesencephalic brain stem of a column of medially placed cells accounts for the well documented arousal effects of stimulations of the reticular formation (see Lindsley, 1961, and Magoun, 1958 for review). Less well known is the fact that such effects are obtained even more rostrally in the diencephalon in a continuation of this neuron system into the hypothalamus.

The contribution of hypothalamic visceromotoric mechanisms to the reflexive aspects of the orienting reaction has been noted by Feldman and Waller (1962). The episodes of fighting and fleeing that are produced by electrical or chemical stimulation to the so-called "defense" region of the hypothalamus can be related to the orienting reaction. Abrahams and Hilton (1958) and Abrahams, Hilton, and Zbrozyna (1964) found that, in attempting to produce a defense response by stimulation of the hypothalamus, at first, a much lower degree of arousal occurred, indicated by pupil dilatation and postural alerting. Only when the level of stimulation was increased and maintained for a few seconds did hissing, snarling, running, and piloerection occur. In the later study, alerting behaviors were measured in greater detail, and during mild stimulation the authors observed changes in pupil dilation, head movement, pricking the ears, respiration, and blood flow. These same changes were also recorded during responses to sim-

ple auditory, visual, or cutaneous stimuli in the absence of hypothalamic stimulation. Since these physiological changes are the same as those observed in all orienting responses, the defense reaction could therefore be considered in part as due to an increase of arousal.

Sensory Systems and the Neuronal Model

There is ample evidence that a host of other brain mechanisms are involved in the orienting reaction. Electrical brain recording, both with macro- and with microelectrodes, shows that initial desynchronization and increase in neuronal discharge occurs in a large number of neural systems. One of the characteristics of familiarization consequent on repetition of a stimulus is the progressive restriction of these sites (John, 1967; John & Killam, 1960). Paradoxically, however, this restriction involves primarily the sites where arousal neurons are located and does not encompass either sensory pathways or the primary sensory projection systems where responses continue to be recorded in the habituated organism. These, therefore, became prime candidates for the locus of comparison of the input stimulus configuration with the neuronal model registered earlier through habituation.

To determine where the detailed comparisons necessary to the Sokolovian model of habituation and orienting might occur, Grandstaff and Pribram (1972) studied the possibility that Thompson's results pertained only to spinal cord and perhaps lower brain stem mechanisms but that at more rostral stations of the sensory projection systems, patterns of change rather than simple decrementing were produced during habituation. Surprisingly, this study showed the Thompson model to hold at the collicular and thalamic levels: electrical responses recorded from the lateral geniculate nucleus as well as those from the retina and superior colliculus (also obtained with microelectrodes by Horn and Hill, 1966; and by Bures and Buresova, 1970) in intact monkeys subjected to repetitive visual stimulation uniformly decremented to repetitive monotonous stimulation, and enhanced responses occurred only when a change in stimulation was initiated. Soko-

lov's model was saved, however, by the results obtained from recordings of the activity of visual cortex. Here, habituation produced no uniform decrementing—some records invariably and reliably showed an incrementing of activity, while others decremented (over each of many days of recording) and still others showed no reliable changes. A *patterned habituation* response occurred, and dishabituation to a change in stimulation disrupted this pattern.

Patterned changes in cortical *rhythmicity* also occur. Thus patterning of electrical brain activity is not limited to abrupt potentials evoked by the stimulus. Earlier studies have been reviewed by J. Mackworth (1969). More recently it was shown that during habituation, changes occur in both latency (increase) and duration (decrease) of desynchronization when baseline activity is in the 10 Hz alpha frequency range (Bagshaw & Benzie, 1968; Grandstaff & Pribram, 1972). More complex frequency analysis of the ongoing cortical activity of limited areas of the visual cortex (Grandstaff & Pribram, 1972) has shown that a decrease in power with habituation in 8–10, 37–50, and 72–80 Hz frequency ranges occurs when the amplitude of the evoked potential decreases. By contrast, in those electrode placements where the amplitude of the evoked response increases, an increase in power in the alpha and other critical ranges is obtained. These patterns of change do not occur at subcortical loci (optic nerve and lateral geniculate nucleus) but are limited to recordings made from the visual cortex, supporting the conclusion that *patterning* of neuroelectric activity, an indication of the construction of a "neuronal model," is primarily the function of a cortical information processing competency.

The occurrence of a patterned response at the cortex suggests that each of the cortical sites can be considered separate information processing channels, which is consonant with the anatomy of these systems. Neisser (1967), Kahneman (1973), and Lindsay (1970) have thoroughly reviewed the behavioral evidence for considering the sensory channels to be primarily parallel processing systems, or at least some form of multiplex-

ing mechanism that relies on an initial simultaneous parallel processing stage, a problem which will be considered in the section on categorizing. Here we first outline the neural systems involved in the arousal and activation that leads to desynchronization of the patterned neuronal model.

Amygdala Circuits

There is evidence for the involvement of the amygdala and related frontal cortical structures in the attentional control of the core brain arousal systems. This evidence delineates two reciprocally acting circuits, one facilitatory and the other inhibitory, both converging on hypothalamic structures related to arousal. Such reciprocal innervation allows sensitive modulation (tuning) of the arousal mechanism.

In pursuing the question of which brain systems are involved in orienting and habituation, we first replicated and showed some of the limits over which the Sokolov results held (Koepke & Pribram, 1966, 1967). When adapting the techniques to the monkey

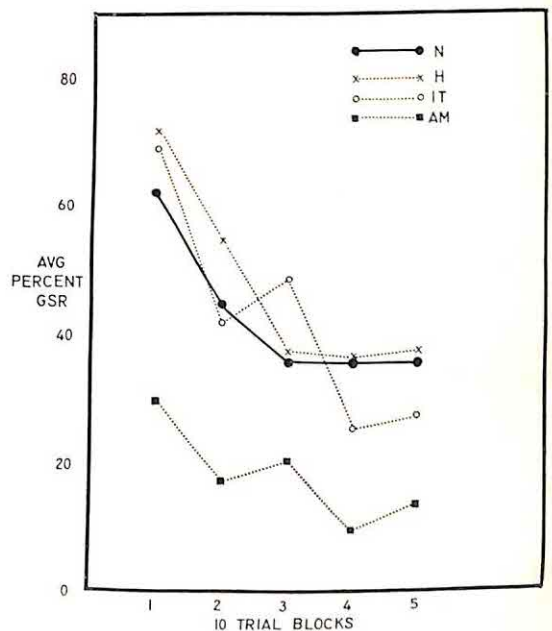


FIGURE 1. Curves for overall percentage of galvanic skin response (GSR) to the first 50 presentations of an irregularly presented 2-sec tone. (Abbreviations: N = unoperated group; H = hippocampectomized group; IT = control group; AM = amygdalectomized group.)

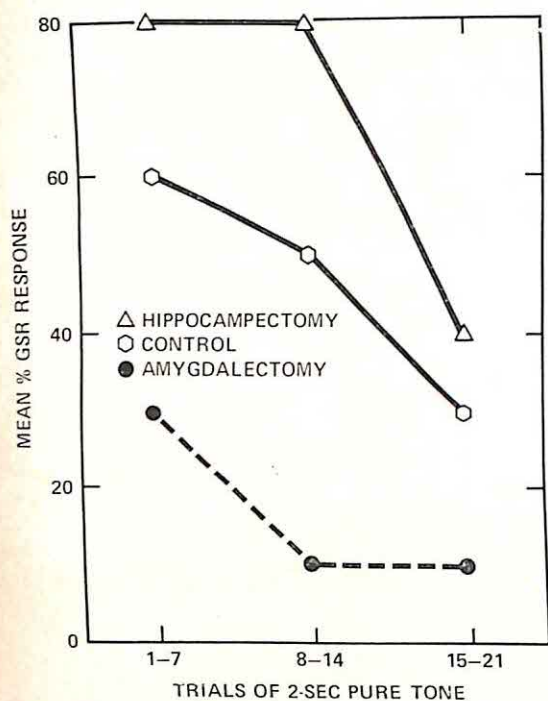


FIGURE 2. Curves for a more detailed analysis of the first 21 trials (split into 7-trial blocks) or percentage of galvanic skin response (GSR) to a 2-sec tone for the amygdalectomized, hippocampectomized, and unoperated monkeys. (See also Table 1.)

we found, contrary to expectation, that lesions of the mediobasal, frontolimbic forebrain, and not of the convexity of the cerebral mantle, had a dramatic effect on measures of orienting (Bagshaw, Kimble, & Pribram, 1965; Kimble, Bagshaw, & Pribram, 1965).

Amygdalectomy results in a failure of behavioral habituation, producing an ever-recurring locomotor orienting reaction (Bagshaw & Benzie, 1968; Schwartzbaum, Wilson, & Morrisette, 1961). Such hyperactivity is also produced by lesions of the orbital surface of the frontal lobe (Mettler & McLardy, 1948; Ruch & Shenkin, 1943; Wall & Davis, 1951) which has been shown to be a part of a larger orbitofrontal-insular-temporal-polar-amygdala system (Kaada, Pribram, & Epstein, 1949). When an analysis of the visceromotor components of the orienting response was carried out on frontal and amygdalectomized monkeys, it was expected that these components would

parallel behavioral orienting. In many instances the opposite was observed.

Resections of dorsolateral frontal cortex *always* abolish the visceromotor components of orienting (Kimble, Bagshaw, & Pribram, 1965; Luria & Homskaya, 1970; Luria, Pribram, & Homskaya, 1964), while amygdala resections produce a bimodal distribution. Electrodermal responses during orienting (Figures 1 and 2; Bagshaw & Benzie, 1968; Bagshaw, Kimble, & Pribram, 1965) are reduced in amplitude and amount for most subjects. However, certain subjects show an increased number and amplitude of response (see Figure 3 and Table 1). Tonic heart rate is paradoxically elevated (see below), but all phasic heart rate responses are abolished (Bagshaw & Benzie, 1968; Pribram, Reitz, McNeil, & Spevack, 1974).

Thus, reciprocal systems appear to be involved—one apparently relates to the dorsolateral frontal cortex since resections of this structure *invariably* eliminate visceral-autonomic orienting responses. The other, op-

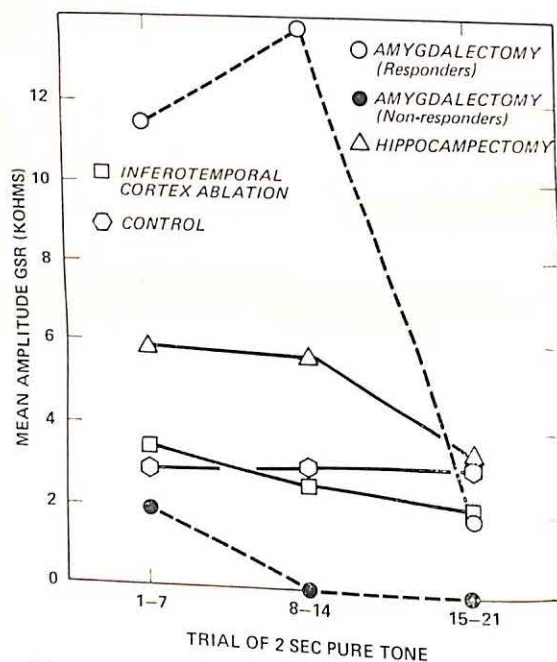


FIGURE 3. Group mean amplitude for galvanic skin response (GSR) to tone on the first 21 trials showing the bimodal distribution of the amygdalectomized monkeys when compared with normal (control) and with two other brain-lesioned groups of monkeys: inferotemporal cortex resection and hippocampectomy.

TABLE 1
AMOUNT AND AMPLITUDE OF GALVANIC SKIN RESPONSES OVER TRIAL BLOCKS FOR SUBJECTS

Group	Response (%)			Amplitude (k Ω)		
	Trials 1-7	Trials 8-14	Trials 15-21	Trials 1-7	Trials 8-14	Trials 15-21
Normal (14)	63	54	43	3.0	2.9	3.2
Hippocampectomy (7)	71*	70*	62	5.9*	5.8*	3.5
Amygdalectomy (6)						
Hyperresponsive (2)	90*	50	40	11.4*	4.3*	1.9
Hypo-responsive (4)	15*	00*	00*	2.0	—	—

Note. Number of subjects in each group appear in parentheses following the group.
* $p < .05$.

posite in function, is most likely related to the orbitofrontal cortex which has been shown to be the rostral pole of an extensive inhibitory pathway (Kaada, Pribram, & Epstein, 1949; Pribram, 1961; Sauerland & Clemente, 1973; Skinner & Lindsley, 1973; Wall & Davis, 1951). These data relate to the early descriptions of the behavior of amygdalectomized animals (Pribram & Bagshaw, 1953) which focused on the fact that while the majority were tame, unresponsive to threat, and nonaggressive, the opposite finding was also occasionally observed (e.g., Rosvold, Mirsky, & Pribram, 1954), and more recent behavioral studies by Ursin and Kaada (1960) using restricted lesions and electrical stimulations have confirmed the suggestion that at least two more or less reciprocal systems can be identified in the amygdala.

Attentional Control of the Arousal System

The presence of reciprocal systems suggests the existence of a locus upon which they converge, a locus which, from anatomical considerations (Pribram & Kruger, 1954), is most likely to be found in the hypothalamic regions of the brain stem. This suggestion is supported by microelectrode evidence at the spinal cord level (Groves & Thompson, 1970) which indicates that the origin of the desynchronizing mechanism leading to orienting may be localized to the medial parts of the spinal cord where the visceromotor outflow originates. The neurons that respond during orienting are likely to have derived their

input from locations near the visceromotor regulating mechanisms. This is borne out both by the results of electrical stimulations of the hypothalamus (noted above), the amygdala (Gastaut, 1953, 1954), and related cortical areas (Wilcott, in press). Thus, not only can the visceromotor components of orienting be sharply distinguished from the somatomotor components, but their role in the orienting mechanism appears to be primary.

The fact that the two amygdala systems are reciprocal, one facilitating, the other inhibiting arousal, suggests a controlling, modulating mechanism and is in accord with evidence on other control functions of the amygdala and related structures. Injections of carbachol into the appropriate hypothalamic site will initiate drinking; such injections into the amygdala have no effect unless the animal is already drinking, in which case the amount of drinking becomes proportional to the amount of carbachol injected in an exquisitely accurate relationship (Russell, Singer, Flanagan, Stone, & Russell, 1968). The frontoamygdala influence can be conceived as a finely tuned determinant controlling visceromotor arousal initiated by the hypothalamic mechanism during orienting. It is as if in the absence of the frontoamygdala systems, the animal would fail to control his drinking behavior: Once started he would drink under circumstances in which others would stop. This is exactly what happens—and more. Both eating and drinking are controlled in this fashion—not only their cessation but also

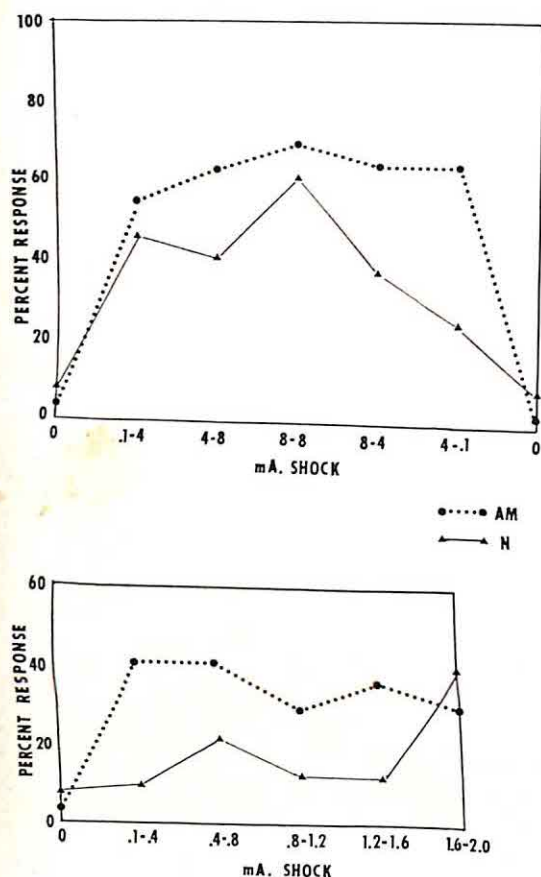


FIGURE 4. Top: Curves of percentage of galvanic skin response generated by three runs of stimuli in ascending and descending intensity (in map) by the amygdectomized (AM) and control (N) groups. Bottom: A finer breakdown of stimulus values from .1 to 1.0 mA, pooled ascending and descending values. (From "Effect of Amygdectomy on Stimulus Threshold of the Monkey" by M. H. Bagshaw and J. D. Pribram, *Experimental Neurology*, 1968, 20, 197-202. Copyright 1968 by Academic Press. Reprinted by permission.)

their initiation (Fuller, Rosvold, & Pribram, 1957).

Sensitivity and Registration

The question arises as to the significance for the orienting reaction of the visceromotor components per se. The absence of visceromotor reactivity during orienting could result simply from lowering the sensitivity of the organism to stimulation. However, resections of frontal cortex (Grueninger, Kimble, Grueninger, & Levine, 1965) or of the amygdala (Bagshaw and Coppock, 1968; McNeil, 1972) do not abolish galvanic skin

responses to movements or to shock. When threshold for shock was measured, frontal lobectomized (Grueninger & Grueninger, 1973) and nonresponsive amygdectomized subjects were found to have, if anything, greater sensitivity (see Figure 4). Specifically, these monkeys frequently produced a maximal number and amplitude of responses to intensities that produced very little or no reaction in normal subjects.

This leaves no doubt that the peripheral response mechanism remains intact. Neither a change in sensory sensitivity nor an altered peripheral response mechanism can account for the effects of frontal and amygdala lesions on the orienting reaction. The critical influence exerted by these forebrain structures appears to relate to arousal induced registration of input effecting changes in the organization of central mechanisms.

Certain changes relate to the finding that despite reactivity to shock, the amygdectomized and frontally lesioned subjects have fewer "spontaneous galvanic skin responses" during the shock sessions, suggesting an altered base level. So far, we have referred to phasic changes in the *initial* period of observation, which often reach asymptote within three to five stimulus presentations. However, while behavioral and some electrocortical responses appear to be normal during orienting, after frontal and amygdala lesions (Bagshaw & Benzie, 1968; Schwartzbaum et al., 1961) the background level of these responses is lower than for controls. For example, ear flicking is practically absent during interstimulus intervals (Bagshaw & Benzie, 1968), and while electromyographic responses occur with normal latency, the amplitude of these responses is considerably reduced (Pribram, Reitz, McNeil, & Spevack, 1974).

Accompanying the reduction in baseline of phasic responses, a striking *tonic* psychophysiological effect followed amygdectomy. This was the finding of a paradoxically elevated basal heart rate (Bagshaw & Benzie, 1968; Pribram et al., 1974). The paradox has been clarified by experimental results obtained by Elliott (Elliott, Bankart, & Light, 1970) in which he found that novel input had little effect on tonic heart rate or

had an effect (deceleratory) opposite to expectations. However, response factors and incentive factors (reinforcing consequences) had strong accelerating effects. Thus, arousal, though it may produce a brief initial phasic acceleration (Obrist, Wood, & Perez-Reyes, 1965), leads more often to tonic heart rate *deceleration*, which is indicative of *activation*, a process detailed in the next section. In short, the monkeys with absent arousal reactions are consistent in showing an elevated tonic heart rate. As Elliott's data indicate, and as will be detailed in the last section of this article, elevated tonic heart rate is manifest when the situation demands "effort" on the part of the organism.

The effects of amygdectomy can be interpreted as follows: Because the specific controls on arousal are removed, behavioral arousal fails to lead to the registration of the situation by altering the neuronal model but results in nonspecific defensive "effort" to cope with the situation. This defense reaction is characterized by an attempt to shut off further input (see Pribram, 1969), and is reflected in an elevated heart rate and other changes indicative of a lack of readiness to respond meaningfully to the input. Thus "effort" is manifest in the absence of readiness. This interpretation is borne out by the results of an experiment in which infant kittens were raised in isolation. When examined at the age of 6 months their visceromotor and endocrine reactivity in orienting experiments was essentially similar to that of amygdectomized subjects: They had not learned to cope with situations and thus showed the "defensive" syndrome suggestive of considerable "effort" (Konrad & Bagshaw, 1970).

Summary

Studies relating brain function and the orienting reaction to sensory input have pointed to the presence of a system of neurons responding to the amount of input to them by maintaining or incrementing their activity. This core system of neurons extends from the spinal cord through the brainstem reticular formation, including hypothalamic sites, and lies in close proximity to those responsible for the engenderment of

visceromotor responses. By way of its diffuse connections, this system is responsible for the more ubiquitous "arousal" responses recorded throughout the brain concomitant with orienting. Forebrain control over this core brain arousal system is exerted by reciprocal facilitatory and inhibitory circuits centered on the amygdala. These circuits control the onset and duration of neural arousal much as they control the onset and duration of visceromotor and appetitive responses.

Our interpretation of the relationship between the lack of visceromotor responses to orienting and the failure to habituate behaviorally has been to suggest that a deficiency is produced in the more ubiquitous central mechanism by which organisms "register" input. When such failure in registration occurs, the organism's nervous system is temporarily swamped by the arousing input and reacts defensively to shut out all further input and thus leads to automatisms. This interpretation fits the clinical picture of the amnesic states (*deja vu* and *jamais vu*) and the automatisms occurring during psychomotor seizures produced by epileptic lesions in the region of the amygdala. There is also considerable congruity between this interpretation and those of Mednick and Schulsinger (1968) and of Venables (Gruzelier & Venables, 1972) in their report of two classes (galvanic skin responders and non-responders) of patients diagnosed as schizophrenic. However, the interpretation also suffers from the difficulties that plague understanding of these clinical syndromes: How do disturbances of registration in immediate awareness influence subsequent retrieval? This will be discussed further in the following sections.

VIGILANT ATTENTION, EFFORT, AND THE CONTROL OF ACTIVATION

The interaction between behaving organisms and their environment is not one-sided. The organism is not just a switchboard for incoming stimulation. Rather, the essence of behaving organisms is that they are spontaneously active, generating changes in the environment often by way of highly programmed, that is, serially ordered, responses

to respond. Yet, under many circumstances the two reactions appear to be yoked: at the cortex by the TNV and in the hypothalamic region by the reciprocity of the "stop" and "go" mechanisms. In such situations they share the function of reflexly coupling input to output, stimulus to response. In the absence of controlled arousal and activation, behaving organisms would be constantly aroused by their movements and moved by arousing inputs. There must be some process that involves both arousal and activation which allows the uncoupling to take place. This process is habituation, which we have seen to be critical in the development of the neuronal model. Action generated inputs (the outcomes of actions, their reinforcing consequences) appear to produce more complexly structured neuronal models than repetitions of simple inputs per se, and reflect the participation of the central motor

systems in generating input. It takes longer to form a habit in, than to habituate to, the same situation. Complex inputs such as repeated exposure to the same musical performances do not readily induce habituation but in such situations a good deal of motor readiness—listening—is also involved.

Having considered the effects of certain limbic lesions on arousal, we need now to review the effects of limbic lesions on habituation and to relate such effects to the activation of vigilant readiness.

Hippocampal (Papez) Circuits

Subjects with bilateral hippocampectomy tend to show a percentage of reactivity and amplitude of the galvanic skin response opposite to that observed in the nonresponding amygdalectomized monkeys (see Figures 2 and 3 and Table 1) in the ordinary orienting paradigm. In addition, the phasic skin response terminates considerably more rapidly in hippocampectomized subjects than in controls (see Table 2 and Figure 5). It appears from this that hippocampectomized monkeys (and amygdalectomized hyperresponders) re-equilibrate more rapidly than normal subjects whose slower galvanic skin response recovery may indicate a more prolonged processing time. As we shall see below, this is consistent with other data that show impaired processing of the disequilibrium produced by a mismatch of input to the neuronal model as a result of hippocampectomy.

A further change is that such subjects show delayed or absent orienting reactions when thoroughly occupied in performing some other task (Crowne & Riddell, 1969; Kimble, Bagshaw, & Pribram, 1965; Rappelson, Isaacson, & Douglas, 1965; Riddell, Rothblat, & Wilson, 1969; Wicklegren & Isaacson, 1963). In short, these animals appear to be abnormally intractable.

But in some situations this appearance of intractability is restricted to the overt responses of the organism, not to orienting per se. Douglas and Pribram (1969) used distractors in a task in which responses were required to each of two successive signals. Hippocampectomized monkeys initially responded much as controls did by overtly

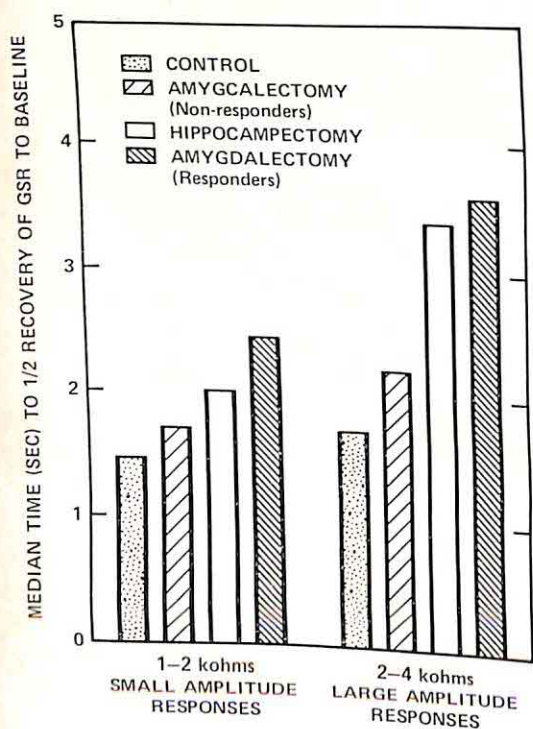


FIGURE 5. Bar graph showing the time taken to attain half recovery to baseline of the visceromotoric perturbation measured as an electrodermal response (galvanic skin response, GSR). (Small, 1-2 Kohms, and large, 2-4 Kohms, perturbations are treated separately since amplitude of response has an obvious effect on recovery time. See also Table 2.)

manipulating the distractors and thus increasing the time between the two required responses. However, in contrast to the behavior of the controls who simultaneously habituated overt manipulation of the distractors and the interresponse time, the hippocampectomized group showed decrementing only of the overt manipulations—their interresponse time failed to habituate at all. In this situation, hippocampectomized monkeys continue to be perceptually distractible while becoming behaviorally habituated and indistractable (Figure 6). This result is identical to that obtained in man with medial temporal lesions: Instrumental behavior can to some considerable extent be shaped by task experience, but verbal reports of the subjective aspects of experience fail to indicate prior acquaintance with the situation (Milner, 1958).

The dissociation between habituation of

TABLE 2
HALF-RECOVERY RATE OF GALVANIC SKIN
RESPONSES FOR SUBJECTS

Group	Half-recovery rate for controlled amplitudes (k Ω /sec)	
	Small 1-2k	Large 2-4k
Normal (14)	.320	.460
Amygdectomy (6)		
Hyperresponsive (2)	.730*	.730*
Hypo-responsive (4)	.250	.460
Hippocampectomy (7)	.580*	.620*

Note. Number of subjects in each group appear in parentheses following the group.

* $p < .05$.

perceptual responses and habituation involving somatomotor performance appears to be part of a more general effect of hippocampal lesions, as it is manifest in other situations

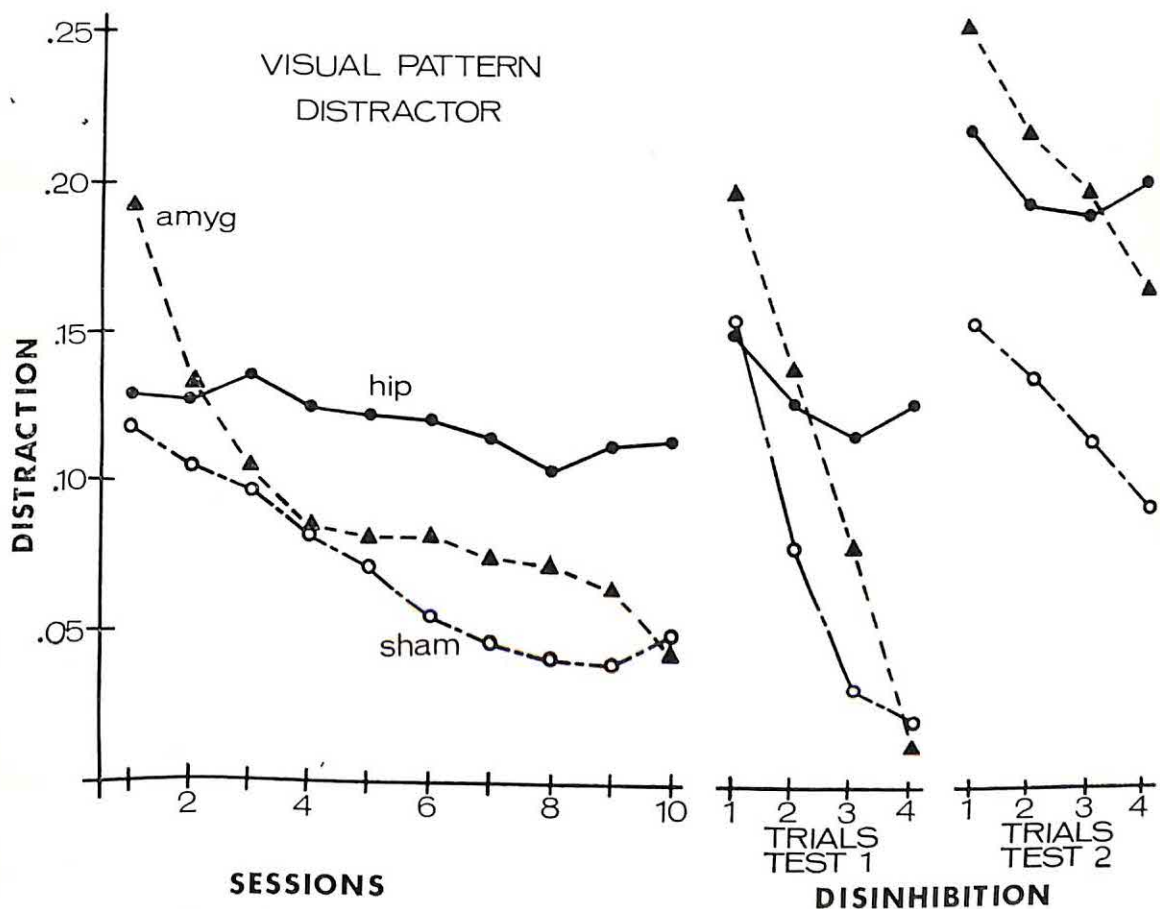


FIGURE 6. Distraction with visual pattern stimulus as distractor. (Abbreviations: sham = sham-operated subjects; amyg = amygdectomized monkeys; hip = hippocampectomized animals.)

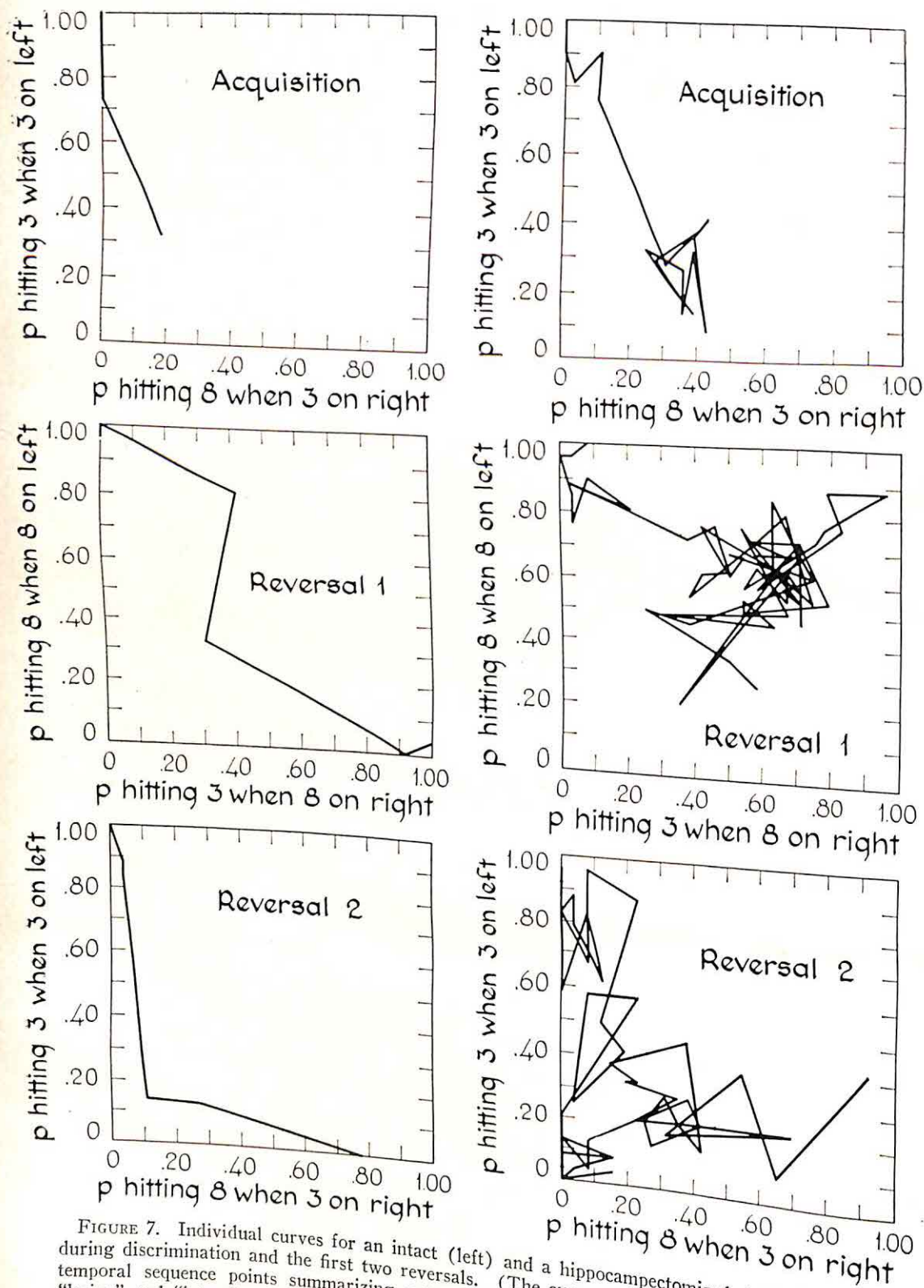


FIGURE 7. Individual curves for an intact (left) and a hippocampectomized (right) monkey during discrimination and the first two reversals. (The curves were constructed by joining in temporal sequence points summarizing pattern-discrimination and position preference of the "lesion" and "intact" monkeys for 50 trial sessions during pattern-discrimination and reversal training.) (From two figures in "A Decisional Analysis of the Effects of Limbic Lesions in Monkeys" by A. Spevack and K. H. Pribram, *Journal of Comparative and Physiological Psychology*, 1973, 82, 211-226. Copyright 1973 by the American Psychological Association, printed by permission.)

in which hippocampectomized monkeys are tested. In a discrimination reversal situation, extinction of previously learned behavior and acquisition of newly correct responses were observed. However, in contrast to their controls, the monkeys with the hippocampal lesions remained at a chance level of performance for an inordinately long time (Pribam, Douglas, & Pribam, 1969). This effect was due to the "capture" of the behavior by a position bias to respond to the 50% schedule of reinforcement—a bias hardly manifest in unoperated monkeys (Figure 7; Spevack & Pribam, 1973). This result suggested that a hierarchy of response sets was operative in the situation such that "observing" responses (indicative of "attention") were relinquished when the probabilities of reinforcement of discriminative stimuli ranged around the chance level.

Taken together, these experimental results suggest that interference with the hippocampal circuit reduces the organism to a state in which the more effort demanding relationships between perception and action, between observing and instrumental responses, and between stimulus and response are relinquished for more primitive relationships in which either input or output captures an aspect of the behavior of the organism without the coordinating intervention of central control operations. The mechanism by which the hippocampal circuit accomplishes this relationship has been elucidated to some extent by recordings of electrical activity from the hippocampus both with micro- and macroelectrodes and by precise electrical stimulations of selected parts of the hippocampal circuit.

The Hippocampus—Arousal and Activation

The evidence that delineates the hippocampal mechanism accounts for the general observation that the *greatest* changes in gross electrical activity observed during habituation are recorded from the core brain (especially mesencephalic reticular and limbic control systems; John, 1967; John & Killam, 1960).

The hippocampus with its three-layered cortex provides the best opportunity for observing unit activity responsible for the

changes observed in gross electrical activity in repetitive situations. Vinogradova (1970) found that *all* neurons of the hippocampus habituate to repetition of a stimulus and dishabituate to any change in the stimulus configuration. But she distinguished two types of neurons, A and I: those activated (30–40%) and those inhibited (60%) by a stimulus, respectively. Habituation occurs by a progressively shortening response in the course of 16–20 repeated presentations. The averaged poststimulus histograms of the two classes are "mirror images of each other."

Some important characteristics of hippocampal units are that (a) the latencies of response to a stimulus do not change; (b) they are of the order of 100–1000 msec even initially; (c) in the ventral hippocampus (the only part present in primates, including man) stimuli must be of a minimum duration of from $\frac{1}{2}$ to 1 sec to produce any noticeable change in the background activity and (d) such facilitation persists up to 1 min after the cessation of the stimulus. These characteristics indicate a necessity for a long period of summation to precede hippocampal facilitation. Vinogradova (1970) interprets her findings as follows:

The duration of reactions in hippocampal neurones shows that the processes continue here long after the information processing is finished in all specific sensory structures, and in primary and secondary areas of the cortex as well. . . . As Gloor (1961) indicated, the quality of sensory information is almost erased in hippocampal neurones. (pp. 114–115)

These results are in accord with proposals previously put forward by Douglas and Pribam (1966). We suggested that the hippocampus constitutes part of an error (mismatch) evaluating mechanism which was conceived to process *only the perturbations* resulting from the mismatch among inputs (including those consequent on responses) (Pribam, 1971). Vinogradova is in agreement, therefore, in suggesting that precise sensory information is *not* involved. Further, she suggests a mechanism by which such processing can be achieved:

The hippocampus exerts a tonic inhibitory influence upon the reticular formation, blocking activatory processes through the tonic discharge of its I-neurones when novelty is absent and registration is

not needed. But when a stimulus which is not registered in the memory system appears, this inhibitory control is blocked (I-neurons become silent), arousal occurs, and the process of registration starts. (p. 114)

Lindsley has recently elaborated the mechanism by which the hippocampally controlled reticular formation can effect these changes in registration. Lindsley (Macadar, Chalupa, & Lindsley, 1974) in keeping with many other recent publications (e.g., Fibiger et al., 1973; Ungerstedt, 1974) has dissociated two systems of neurons that influence the hippocampal circuit. One system originates in the median raphe and associated structures of the mesencephalic reticular formation, the other originates more laterally in the locus ceruleus and other portions of the periaqueductal grey. We have come to know these two systems in the present review as a serotonergic (indole amine) "stop" mechanism associated with arousal, and a dopaminergic and norepinephrinergic catechol amine "go" mechanism associated with readiness and activation. Lindsley's findings were obtained by electrical stimulations of the appropriate structures in the mesencephalic reticular formation. Such stimulations of the stop mechanisms produced hippocampal desynchronization and at the same time a synchronization of the amygdala circuits. Taken together with Vinogradova's evidence, this suggests a reciprocal process by which the controls on arousal are maintained as long as hippocampal inhibition of the reticular formation is in progress—much as Vinogradova suggests. Only when a mismatch from the neuronal model is signaled to the reticular formation does this inhibitory control become loosened producing hippocampal desynchronization—and concomitant relaxation, synchronization, of the arousal functions of the amygdala circuits. Lindsley has found that often, though not always, such hippocampal desynchronization is accompanied by desynchronization of the sensorimotor projection systems, suggesting that "registration," an alteration of the neuronal model of the cortical representation, is occurring. Note that in this formulation the term "registration" refers to a change in the neuronal model, thus a "registration in memory," a

process that, as we shall see, requires effort. This use of the term registration must be distinguished from "registration in awareness" which, as noted earlier, is disrupted by interference with the amygdala circuits.

Theta Rhythms and Intent

The second mechanism discerned by the Lindsley studies is the catechol amine "go" system which has been the focus of this section. This mechanism makes possible the "What is to be done" reaction, the processing of response-produced inputs. When electrically stimulated, the mesencephalic portions of the "go" mechanism initiate hippocampal rhythmic activity in the theta range of frequencies. Early studies (Green & Arduini, 1954) had uncovered the paradox that the desynchronization of the electroencephalogram recorded from the brain's convexity during "activation" was accompanied by synchronization in the recordings obtained from the hippocampus. Though such synchronization is not as obvious in records obtained in monkey and man, computer analysis has shown it to occur and that it can also be studied in the primate (Crowne, Konow, Drake, & Pribram, 1972). This synchronous rhythm is in the theta range (4–8 Hz) and has become the focus of a long series of studies.

That theta frequencies are especially prominent in records made from the hippocampus was noted by Jung and Kornmuller (1938). Later a series of studies (Green & Arduini, 1954; Grastyan, 1959; Grastyan, Lissak, Madarasz, & Donhoffer, 1959) described the occurrence of hippocampal theta to the orienting stage in a conditioning situation. Still later, however, the occurrence of hippocampal theta was related instead to intended movement (Dalton & Black, 1968; Black & Young, 1972; Black, Young, & Batenchuk, 1970): intended rather than overt movement because the theta rhythm occurs in completely curarized preparations who have been trained in the uncurarized state to lever press. These conclusions were in part anticipated by another series of experiments conducted by Vanderwolf and his associates (Bland & Vanderwolf, 1972a, 1972b; Vanderwolf, 1969, 1971; Whishaw, Bland &

Vanderwolf, 1972). Rats were observed while moving freely in the test situation and hippocampal electrical activity was recorded continuously. Theta activity occurred almost exclusively when the rats were making "voluntary" movements, defined as acts or response sequences characterized by flexibility and modifiability.

Adey and his group (Adey, 1970; Adey, Dunlop, & Hendrix, 1960; Elazar & Adey, 1967; Radulovacki & Adey, 1965) took the investigation of behavioral effects on hippocampal theta rhythms a step further by showing that a shift (4-6 Hz) in power within the theta frequency range occurs in cats from the prestimulus period, through stimulus presentation, to the correct response. Bremner (1970) has further analyzed the changes in theta activity along three dimensions: (a) an increase or decrease in the total amount of power in the theta range; (b) a narrowing or broadening of the range of energy distribution around the peak frequency; (c) the location of that peak frequency in the electroencephalogram spectrum. Change in total power of theta depends on the visceromotor arousal (increase) versus somatomotor readiness (decrease) distinction delineated here, and the changes in contour and their distribution correspond to the categorizing-reasoning distinction to be pursued in the latter parts of this review.

Specifically, Bremner has used the habituated organism (rat and man) as a baseline preparation and then studied manipulations which gave rise to orienting, simple conditioning, discrimination, and discrimination reversal. He found that the power (amount of energy) of hippocampal theta increased as a function of the stimulus which produces orienting and arousal but decreased in the anticipatory interval preparatory to response in the conditioning situation. Narrowing reflects discrimination, and the location of the peak is shifted by discrimination reversal procedures, findings especially relevant to the results of the behavioral experiments reviewed above (Figure 8).

Thus Bremner finds that the power of theta produced by a distracting stimulus becomes attenuated when an organism is already preoccupied in the performance of a

task. Further, during discrimination reversal, despite the fact that the now positive stimulus elicits enhanced power in the theta frequency range within two or three trials, extinction of the prior correct stimulus shows a long time course and the dominant frequency may appear to shift in one direction or another (Gray, 1971). These data also suggest an explanation for the finding detailed above (Pribram et al., 1969) that hippocampectomized monkeys are little different from normal subjects in reversal learning except that they remain fixed at the chance level for an incredibly long period.

Summary

By way of review, we note that there exists evidence for the organization of a central representation of input, the construction of a (cortical) neuronal model. Changes in this neuronal model are controlled by two subcortical systems: One is located in core portions of the neuraxis and contains neurons that increment to or monitor input, while another more laterally placed, contains neurons which rapidly decrement when they are repeatedly stimulated. In addition, we have delineated three brain mechanisms in the rostral portions of these control systems: (a) The first centers on the amygdala, which regulates the monitoring or "arousal" neurons and becomes organized into a "stop" or reequilibrating mechanism. (b) A second is centered on the basal ganglia and involves the activation of "go" mechanisms—expectancies (perceptual) and readiness (motor). (c) Finally, a third mechanism comprising

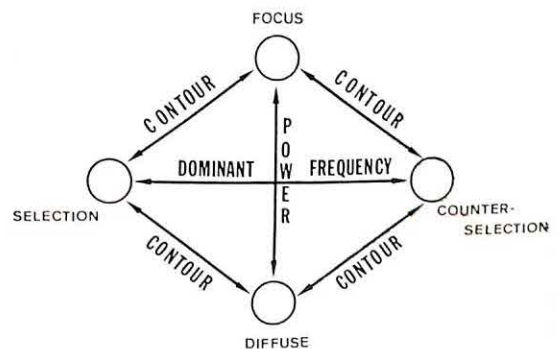


FIGURE 8. Diagram of the relationship among dimensions of hippocampal theta activity and behavior.

the hippocampal circuit has been identified, which uncouples the stimulus from the response by coordinating the amygdala and basal ganglia mechanisms so that appropriate changes in the central representation can occur, a process that entails *effort*.

As the amygdala is often considered to be one of the basal ganglia, albeit related specifically to the olfactory and visceral systems, brain structures similar in morphology, though different in their connections, control arousal and activation. Their coordination in effecting change in central representation, on the other hand, devolves on a circuit—the hippocampus—very different in morphology, and more akin to that characterizing the cerebellum, another mechanism notable for coordinating complex sequential processes.

Control by the amygdala is achieved by a reciprocal mechanism: a facilitatory lateral-frontal-amygdala-lateral-hypothalamic and an inhibitory orbito-frontal-amygdala-medial-hypothalamic circuit have been identified. Both circuits operate on the monitoring (e.g., satiety) neurons which in the brainstem appear to be serotonergic, that is, indole amine neurochemically. We suggest that the controls modulate the intensity and duration of the reaction initiated and thus modulate the organism's arousal, a "What is it?" reaction. This suggestion can be readily tested by electrically stimulating amygdala (and frontal) locations (much as was done by the Lindsley, 1961, and Clemente et al., 1964, groups) and recording with microelectrodes (as in Richard Thompson's experiments, cited in Groves and Thompson, 1970) to determine the effects on threshold and duration of the reaction initiated by the monitoring mechanism in the appropriate spinal or brainstem (reticular, hypothalamic) neurons.

Control over the "go" mechanisms responsible for the preparatory *activation* of brain systems appears to center on the basal ganglia. Activation of any particular part of the brain is signaled by the gradual development of a negative potential which becomes abruptly positive when the perceptual or motor response is executed. There is evidence that this slowly developing negativity per se is primarily of local origin. The mechanism and pathways involved in the

controlling function are apparently dopaminergic, part of a more extensive catecholamine system.

Hippocampal control over the organization of the relationship between arousal and activation leads to changes in central representations which may be conceived as changes of state, set, or "attitude." Such change, as illustrated in the final section of this review, demands "effort." Hippocampal function involves the notorious Papez circuit, largely invoked in discussions of brain function in emotion, conceived within a visceral theory framework (MacLean, 1949). It is an interesting historical note that Papez put forward his proposal within the context of Nina Bull's attitude theory of emotion (1951) and thus comes close to the views of the function of this circuit as proposed here.

What remains to be systematically investigated is the relationship between hippocampal electrical activity and the variety of electrical manifestations generated in the remainder of the brain. For instance, both the CNV with its concomitant cortical desynchronization and hippocampal theta rhythms have been involved in the intended execution of perceptual and motor responses. Usually study of dc shifts of the CNV type are prevented by the apparatus used to study the electroencephalogram frequencies, and the CNV has not been studied in relation to a frequency analysis of the tonic, ongoing electrical activity of the brain. One exception, already mentioned (Grandstaff & Pribram, 1972) showed that power in the theta range increases dramatically at the *end* of the CNV when a sharp positivity occurs. The remaining question concerns the correlations between the locations and conditions effecting decrease in power in the theta range in the hippocampus (as found by Bremner, 1970) and the buildup or terminal deflection of a CNV. The causal relation between the two types of electrical manifestation can be investigated by examining the CNV in subjects with hippocampal (and other Papez circuit) lesions.

The division of reflex and vigilant attention into arousal and activation components receives substantial support from an analysis

of the components of the event-related gross potentials evoked in the auditory system by Picton and Hillyard (1974). Their succinct summary speaks for itself:

In conclusion we should like to propose a tentative synthesis of all such data into an elementary model for the physiological basis of human auditory attention. The stability of the early [initial] components of the evoked response would seem to indicate that auditory information is analyzed in the lemniscal or primary auditory system in much the same manner regardless of whether it is attended or not. A secondary auditory system, imperfectly defined but probably comprising reticular formation, medial thalamus, and (frontal) association cortex, is involved when further evaluation of the significance of this auditory information is required. *Stimulus set* directs the preferential input to this secondary system from that part of the lemniscal system involved in processing the attended sensory channel. *Stimulus set* is evidenced, therefore, by the increased N_1 - P_2 [large early components of the evoked potential] response of the frontal association cortex. The secondary system functions to compare input from the primary auditory system with memory models or templates of expected or significant stimulus alternatives. Once a relevant or significant signal has thereby been recognized, this decision is followed by appropriate *perceptomotor* sequelae. These sequelae, reflecting the contingencies of a *response set* mode of attention are associated with the generation of the widespread P_3 [late components of the evoked potential] complex recorded from the scalp. (p. 199)

In concluding this part of the review, we want to emphasize for behaviorally oriented psychologists the data we believe to be important to their interests.

Three neurally distinct and separate attentional systems—arousal, activation, and effort—operate upon the information processing mechanism. The presumed operation of these control systems is perhaps best illustrated as follows: The orienting reaction involves arousal but no activation; vigilant readiness involves activation but no arousal; the defense reaction involves both arousal and activation; when neither arousal nor activation is present, behavior is automatic, that is, stimulus-response contingencies are direct without the intervention of any of the control mechanisms of attention. Such automatisms, when they occur abnormally, are diagnostic of lesions affecting the medial surface of the temporal lobe, the site of two of the three control systems, and have also been

experimentally produced by extensive lesions of the basal ganglia (Mettler & Mettler, 1942).

Clearly, in the intact organism these control systems continuously interact with each other as well as with the information processing systems where the representational mechanisms become constituted. Figure 9 may help to summarize in highly oversimplified form the mechanisms detailed thus far, and it suggests the relationship to them of the various higher level cortical controls which are the focus of the discussion in the remaining parts of this review.

SELECTIVE AND VOLUNTARY ATTENTION

Whenever a problem demands a response to an invariant combination of input events (including those which are response produced) the task is usually referred to as a discrimination. The invariant combination must be detected, categorized, and selectively attended to make the appropriate response. This selective attending operates on the processing of information.

Another sort of problem solving occurs when the combination of input events is variable but computable. We will refer to these tasks as reasoning problems. In nonhuman primate experiments, the delayed response and delayed alternation are examples of such tasks; in human experimentation arithmetic calculations and other complex problem-solving exercises have been used.

Reasoning demands the uncoupling of attention from the immediate input variables—the attention involved is *voluntary*, therefore, in the sense that it is initiated by the organism rather than by some input event. Such intentional voluntary processes have recently been studied extensively and are no longer the complete mystery they once were (McFarland, 1971; Miller et al., 1960; Pribram, 1962).

The steadiness of the visual world during voluntary eye movement (as noted by Helmholtz, 1924) is not achieved through feedback from eye muscles but through a parallel processing feedforward biasing (Pribram, 1971, chap. 5) or open loop (McFarland, 1971) mechanism. Detailed experimental analyses by Merton of both eye movement

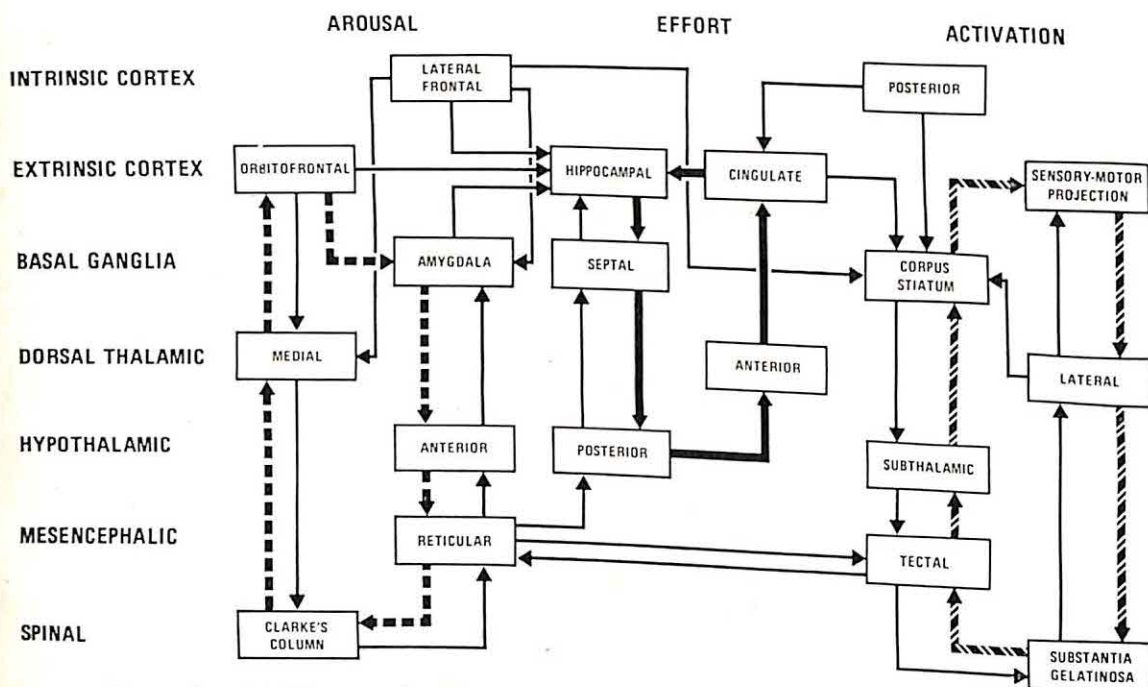


FIGURE 9. A highly oversimplified diagram of the connections involved in the arousal (amygdala), activation (basal ganglia), and effort (hippocampal) circuits.

(Brindley & Merton, 1960) and voluntary movement involving the pyramidal and gamma motoneurons (Merton, 1953) have been reviewed elsewhere (Pribram, 1971, chap. 13) and mark a beginning toward the scientific understanding of the term "voluntary."

An operational definition of selective and of voluntary attention, provided by the discrimination and reasoning tasks described above, will be used to illustrate how categorizing and reasoning involve the three attentional processes discussed earlier. However, we must begin with the delineation of two additional higher order (cortical) neural systems which control central competency by virtue of the organization imposed on incoming information.

Central Competency

In living systems, an arousing stimulus often *increases* the uncertainty of the organism by its novelty. This effect of input information is contrary to that obtained in nonliving communication systems where the information conveyed always *reduces* uncertainty. The difference between living and nonliving systems can be conceptualized in

terms of the channel over which the communication takes place. In nonliving communication systems the channel is akin to a sensorimotor channel which is fixed in capacity and does not alter with the communication. Living systems (and also computers) have the capability of memory which alters the *competence* with which they process information (Pribram, 1971, chaps. 14 & 16). This is produced by altering channel redundancy and superficially resembles a change in the number of channels with fixed capacity. The increase in competence is the result of increasing the complexity of the neuronal model, an encoding process described as "chunking" the information (Miller, 1956; Simon, 1974). This and similar mechanisms in human information processing effect a change in central processing very different from that produced by a simple increase in the number of fixed-capacity channels available.

The evidence that information processing competency can be changed in living organisms comes from a variety of problem-solving situations. Kahneman (1973), in reviewing several such studies from the psychophysiological literature, suggests that "arousal" is

in fact an indicator of a change in capacity—"the allocation of spare capacity"—much as this is changed in nonliving systems by increasing the number of channels available. He also goes on to equate "arousal" and "capacity" with "effort" and "attention" in a global fashion. We hope to have shown already that arousal is indicative of but one sort of attention, and we will now show that effort is involved only when the situation demands the regulation of arousal and activation to produce a change in information processing competency.

The way in which competency is controlled by brain systems in the living primate is demonstrated by the finding that removal of the area of the brain usually called sensory or posterior intrinsic or "association" cortex reduces the sampling of novel alternatives. The opposite effect is obtained when the lateral frontal cortex is resected. It was noted earlier that removal of this same frontal cortex led to an increase in behavioral orienting and an abolition of the visceromotoric components of orienting. There thus appear to be opposite effects (posterior and frontal) on the number of alternatives sampled in a situation. This was interpreted to indicate a dual control mechanism determining the ability to sample (Pribram, 1960a).

Supportive behavioral evidence came from an experiment by Butter (1968, 1969) in which he investigated the number of features usually attended by monkeys while discriminating between two cues. He did this by eliminating each feature in turn in various combinations. He found that resection of the same brain region (the posterior cortex) that produced a restriction in the number of alternatives sampled also produced a restriction in the number of features used to make the discrimination.

Electrophysiological evidence has been obtained that the posterior and frontal cortex contribute opposing controls on sensory channels. This evidence is based on changes produced in the recovery cycles of the system (the speed with which the system recovers to its full capacity after a sudden, intense stimulus) and the alterations produced

in the shape of visual receptive fields (Spinelli & Pribram, 1966, 1967).

These changes in sensory channels were, however, *not* attributed to a simple change in the *number* of channels of fixed capacity, as the effects of surgical resection have shown that as little as a few percent of an anatomically defined sensory channel is sufficient for ordinary discrimination learning, performance, and transfer (Chow, 1970; Galambos, Norton, & Frommer, 1967; Lashley, 1929). The remainder of any input channel appears to be redundant, spare channel capacity, under most circumstances. We therefore interpreted (Pribram, 1967) the results on the control of input channels by posterior and frontal cortex as influencing *redundancy*, *not sensory capacity* in the usual information theoretic sense. Specifically, we suggested that the input systems acted as channels in which spatial and temporal multiplexing could occur, a suggestion similar to that put forward by Lindsay (1970).

We would like to suggest, therefore, on the basis of the data reviewed above, that Kahneman's (1973) concept that arousal involves an increase in the number of sensory channels available be generalized to include constraints involving the redundancy characteristics, the competency, of that capacity. Kahneman's discussion approaches such a generalization when he speaks of changes in "structural connections between components." In technical language, such changes in competency would be reflected in changes in the equivocation of the channel (defined as the sum of noise and redundancy). Competency is the reciprocal of equivocation. Effort can then be defined as the measure of the attention "paid" to increase or maintain efficiency by reducing equivocation, that is, enhancing competency.

Internal and External Constraints

Garner (1962), in his analysis of the structure of redundancy, has shown that the total amount of constraint operating in any system of variables can be divided into internal and external components. Internal constraints refer to the relationships among the systems of variables under consideration, while external constraints refer to the rela-

tionship between these variables and some external referent system of variables. In our neurophysiological experiments we consider the constraints that describe the central operation of the channel as internal and the constraints that refer to operations on the environmental situation which control its sensory input as external. In addition, we have found it important to distinguish between temporal (repetition of the use of the channel or variable over time) and spatial (replication of the variable over space) redundancy for each of Garner's categories.

Specifically, we have suggested (Pribram, 1967) that when the frontal system becomes involved in the orienting reaction, the internal redundancy in the input channels is increased so that all of the information being simultaneously processed becomes chunked into one unit. By contrast, when the posterior cortex becomes involved in the attentional process, internal redundancy in the input channels is decreased, separating the bits of information in each channel from each other. This is concomitant with enhancement of external redundancy which, according to Garner's findings, enhances the ability to make discriminations, that is, to categorize input.

In short, we suggest that the controls on arousal and activation operate on the mechanisms of redundancy, on the constraints operating within and between channels rather than on the information being processed. These constraints involve the neuronal model and may be conceived of as operating on *memory* rather than on input information. Another way of stating this is to say that the controls operate on the context in which the informational content is processed.

A good deal of additional evidence can be cited to show that competency rather than sensory channel capacity per se is controlled by the attentional systems discussed here. For instance, the studies of Anderson and Fitts (1958) cited by Garner (1962) show that as much as 17 bits of sensory information can be simultaneously processed. The work of Lindsay (1970) which demonstrates the relationship between sensory discriminability (difficulty in distinguishing between inputs) and central processing competency

has already been mentioned. Pribram, Lim, Poppen, and Bagshaw (1966) and Mishkin and Pribram (1955), using various forms of the delayed alternation tasks, attributed the differential effects obtained after resections of the two reciprocal frontoamygdala systems as due to selective alterations in the structure of internal redundancy (spatial and temporal, respectively) of the remaining processing competency. Further, Pribram and Tubbs (1967) have shown that when the delayed alternation task, the nemesis of monkeys with frontal lobe resections, is externally parsed or chunked by making the intertrial intervals asymmetric, the deficit is completely overcome. Similarly, M. Wilson (1968) analyzed the trade-off between tasks involving external temporal and spatial redundancy in reciprocal mechanisms (anterior and posterior infero-temporal cortex) which have been delineated within the posterior system.

Thus, both Kahneman's (1973, pp. 8, 9, 15) analysis and ours attribute the control of arousal and selective attention to alterations in information processing channels, not the direct control on information and uncertainty per se. We differ in that Kahneman focuses on the problem of increasing the number of channels of fixed capacity—the "allocation of spare capacity"—while we emphasize the broader issue of competency, defined by any constraints operating on the structure of channel redundancy. We also differ in separating activation from arousal and in that we do not identify attention, arousal, activation, and effort as different names for the same process. Finally, we specify that effort accompanies only those attentional processes that result in a change in the representational organization of the information processing mechanism. Part of the mechanism detailing how and what effect is expended during attention has been revealed by studies measuring peripheral autonomic and somatic changes.

The Psychophysiology of Categorizing and Reasoning

The important viscerosomatic components in orienting have already been outlined. In studies in which perceptual and motor readiness are invoked, the involvement of the

autonomic nervous system becomes more specific due to activation (basal ganglia) of the central representation (sensorimotor cortex) of the somatomotor mechanism, where phasic and tonic mechanisms become separately involved. Further, the sympathetic system more often mirrors stimulus parameters such as novelty and complexity, whereas the parasympathetic division which controls heart rate reflects the parameters of the somatomotor readiness and response mechanism—for example, the incentives, response biases developed on the basis of reinforcing consequences of behavior. This distinction has been clearly demonstrated in a study by Elliott (1969; Elliott et al., 1970) as we have reviewed, by Dahl and Spence (1971), and by Hare (1972).

Experimenters, using autonomic measures during discrimination tasks in which input must be categorized in order to make an appropriate response, appear to use the term "arousal" as actually defined by the *amount* of categorizing required, and studies consistently have shown that sympathetically innervated phasic responses distinguish between variations in task largely on the basis of their collative, informational variables: their novelty, surprisingness, or complexity (Blaylock, 1972; Hare, 1972; Kilpatrick, 1972; Lacey & Lacey, 1970; McGuinness, 1973). A typical experiment showing this effect was that of Hare (1972) in which subjects viewing slides without making discriminations between them were less aroused (lower skin conductance) than those who had to categorize, and this was only an initial effect; phasic skin conductance responses rapidly returned to baseline (habituated) over trials. Difficulty due to changes required in response biases were reflected in tonic skin conductance but most precisely in the heart rate changes.

The absence of a precise relationship between galvanic skin response and response variables has not yet been demonstrated which may be because the galvanic skin response has no known metabolic function (Venables & Christie, 1973). Current data show that the involvement of somatomotor responses enabled by readiness, as well as tonic cardiovascular changes, provide the

clues to unraveling the controversial and apparently disparate views of the psychophysiological mechanism involved in attention.

Readiness and Execution

Two prevailing theories concerning the meaning of cardiovascular processes during attention appear diametrically opposed. On one hand there is the baroreceptor feedback theory of Lacey (Lacey, 1967; Lacey & Lacey, 1970) which, simplified, states that heart rate will directly affect stimulus intake or rejection through a process of feedback to the bulbar inhibitory centers in the brainstem, reducing arousal of the central nervous system (Bonvallet & Allen, 1963). On the other, Obrist and his colleagues (Obrist, Webb, Sutterer, & Howard, 1970a, 1970b) propose that heart rate is regulated by the motor demands of the organism and that these demands are controlled centrally. Findings by Jennings, Averill, Opton, and Lazarus (1971) have challenged the adequacy of either of these explanations and have produced results in line with an "amount of attention" hypothesis. They argue that as demands on attention increase, heart rate will *fall* in proportion to the complexity of the categorization demanded by the task, unless metabolic activity is engaged. This is somewhat in line with the view of this article, but their suggestions are still incomplete. As they note, this hypothesis does not explain the functional significance of cardiac deceleration during attention. Nor too, we might add, does it explain why heart rate always *increases* during reasoning.

To resolve these conceptual difficulties, the physiological processes will be examined in more detail. The assumption of the majority of theories is that there is a unitary relationship between cardiovascular and somatic processes whereas evidence shows that the situation is far more complex.

Any registered orienting response produces a dual sympathetic-parasympathetic effect. The cardiovascular reflex in orienting results in sympathetically controlled and cholinergically mediated gross blood flow shifts which are sometimes accompanied by an initial phasic heart rate acceleration. This will vary in magnitude with stimulus in-

tensity (Graham & Clifton, 1966) and an initial respiratory block will often combine to elevate the acceleratory effect (Jennings et al., 1971; Petro, Holland, & Bouman, 1970; Wood & Obrist, 1964), although the effect of intensity is not entirely predictable (Bull & Lang, 1972). We now know that this is due to parasympathetic inhibition and that all attentional (as opposed to movement-produced) heart rate changes are under vagal regulation (Eckberg, Fletcher, & Braunwald, 1972; Forsyth, 1970; Obrist, Howard, Lawler, Sutterer, Smithson, & Martin, 1972).

The initial phase is followed immediately by a heart rate change in the direction of deceleration, which is due to the vagus restabilizing the system. If the prepared somatic systems become mobilized to initiate an appropriate consummatory or defense pattern—eating, drinking, withdrawal, flight, or attack, (Abrahams et al., 1964)—heart rate will then come under the control of the somatomotor system and a tonic increase in rate occurs to meet the demands of the activity. The processes occur in fixed order: stimulus → blood flow shift → heart rate acceleration → heart rate deceleration, until the motor response determines more tonic heart rate acceleration and the system ultimately restabilizes.

During a vigilance or categorization task, the time course of the responses is somewhat different (Lacey & Lacey, 1970). Now the vagal restabilization phase is extended as part of a tonic activation—the vigilant readiness process. Lewis and Wilson (1970), for instance, examined cardiac responses to a picture-matching task in young children. In all children, a marked deceleration occurred which lasted until the choice was made. The most interesting finding was that the greatest deceleration was related to *correct* responses—if a subject was asked to “guess again,” heart rate did not return to baseline until the final choice. Correct responses occurred with longer response times and greater cardiac deceleration, which supports the suggestion that the vigilant readiness phase of the control processes has been extended.

Correlation to IQ demonstrated that there

was a relationship to both deceleration and error score for girls only. The sex difference finding could possibly be due to the greater amount of movement found in boys which would cause a heart rate acceleration and less attention to the task. This explanation is made more tenable by the recent finding of Obrist, Howard, Hennis, Murell (in press) that children's reaction times are slower in proportion to the amount of uncontrolled movement they produce. The relationship between motor control and reaction time was linear with age. As reaction times increased, motor responses and heart rate decreased. We therefore suggest, in line with Obrist et al. (1970a, 1970b), that when activation must be maintained, heart rate deceleration occurs as an accompaniment of somatomotor readiness.

This then partially accounts for the sequence of physiological events in a certain type of problem solving. When the organism prolongs attention to the complexity of input in a categorizing discrimination situation, he reduces extraneous noise by eliminating random movements. Heart rate then adjusts to the reduced demands of the system reflecting characteristics of the somatomotor system in which overt movements regularly precede blood flow changes and heart rate acceleration. Petro et al. (1970) found that after a voluntary contraction of the biceps muscle, heart rate increased with a latency of approximately 500 msec. However, when reasoning is initiated with its emphasis on which response to produce when, and the actual trial-and-error making of responses, the cardiovascular and somatomotor response relationship is the reverse.

Muscular Effort and Problem Solving

The locus of the demand for effort is different in categorizing and in reasoning, and this raises the question of the relationship between the various forms of attention and effort and also of the nature of effort itself. Could it be that the muscular contractions involved in making responses are totally responsible for the effort necessary to problem solving? A change in actual movement such as during categorizing and reasoning would obviously be accompanied by a decrease

(categorizing) or increase (reasoning) in isotonic contraction, but, in addition, activation of the readiness mechanism may also change a special sort of metabolism, usually thought to be largely anaerobic, which has been shown to be involved during concentration and problem solving and involves isometric contraction.

An understanding of this special metabolic process during attention has been provided by Berdina, Kolenko, Kotz, Kuzetzov, Rodinov, Savtchenko, & Thorevsky (1972). Until this study there was no functional explanation for the finding that sympathetically innervated blood flow shifts actually reduce oxygen uptake by the muscle while somatic activation increases it (Rosell & Uvnäs, 1962). The study of Berdina et al. suggests the possibility that this special sort of metabolism of muscle fibers is due to cholinergic mediators. Berdina et al. were interested in the effect of problem solving on blood flow and muscle contraction. Until their study, it had been extremely difficult to establish any precise relationship between electromyogram recording and problem solving due to the small amount of amplification produced by the technique, which often results in unreliable values (Jennings et al., 1971; Lader, 1965; Matthews & Lader, 1971). To avoid this difficulty they introduced a condition of partial muscular contraction: Subjects had to grip a hand dynamometer which they were pretrained to maintain at 30% of their maximum grip. Then, during the problem-solving experiment, subjects were asked to grip the dynamometer for as long as possible.

Both contraction alone and problem solving (arithmetic) alone caused significant blood flow changes. However, a combination of contraction plus arithmetic produced no greater changes in blood flow than either condition alone. On the other hand, the contraction *duration* was significantly increased by arithmetic from 181 to 235 seconds ($p < .001$). To disentangle voluntary effects from the purely reflex effects, they induced a forearm contraction by electrical stimulation. The results were the same even when the subject had no voluntary control over his muscles. An injection of atropine significantly decreased blood flow during arith-

metic and during contraction plus arithmetic but not during contraction alone. Atropine also reduced contraction duration during the combined task showing that the increased contraction during problem solving was *not* a somatic effect. Since small amounts of circulating atropine do not affect the neuromuscular junction, it was concluded that some form of sympathetically controlled metabolism was causing muscle force to increase, making muscles "work" under the special conditions of isometric contraction. Apart from other considerations, this experiment demonstrates conclusively that muscular effects do occur during problem-solving tasks and that they involve gross changes in tonicity which cannot always be observed by electromyogram recording. Equally relevant is the fact that as heart rate correlates so precisely with blood flow changes (Matthews & Lader, 1971), an increase in blood flow during problem solving automatically implies a corresponding increase in heart rate. The Berdina experiment thus explains the functional significance of the heart rate changes during categorizing and reasoning: Maintaining attention while problem solving is effortful and involves "hard work" and hard work is accompanied by metabolic changes. It does appear from this study that the outlines of the metabolic processes operating during intense concentration have been established.

In summary, the discovery that two types of muscle activation (isotonic and isometric) and two metabolic functions (aerobic and anaerobic) are involved in both voluntary attention and action makes it easier to understand why some confusion has arisen in the study of the psychophysiology of attention. Returning briefly to the Lacey hypothesis, physiological responding during stimulus intake will depend entirely upon how a task is constructed and whether it is primarily a categorizing or reasoning problem. We agree with Lacey that during a state of behavior characterized by heart rate deceleration, we are most open to environmental cues (Lacey & Lacey, 1970) but not necessarily for the reasons he suggests (see also critique by Hahn, 1973). Lacey's theory also requires acceleration when decisions are

made. How then could one interpret what is occurring in acquiring a categorizing or motor skill? When the coordination of arousal and activation is demanded, heart rate acceleration reflects the amount of effort involved in registering, adjusting, and changing the central representation to the requirements of the task. In accord with Lacey's assumptions, intake of information must alternate with concentration. A process ensues that must constantly shift between relevant stimuli in order to overcome the rapid decrementing of the system (just as continuous eye movements overcome the rapid adaptation of retinal receptors). When one plays tennis, for example, heart rate is elevated, but the player must flexibly attend (be aroused by) such cues as the angle and velocity of the approaching ball, the ground lines, the height of the net, and the position and angle of the opponent's body and his racket. As well, the intricacy of the series of highly refined changes in response and the transformational calculations based on this evidence could also elevate heart rate, even without the behavioral activity of playing tennis occurring (e.g., in a spectator). This is brought out clearly in a study by Johansson and Frankenhaeuser (1973) in which high heart rate accompanied the complex transformations required during an intake task.

By contrast, in reasoning tasks, activation precedes arousal. For instance, in playing a game of chess or engaging in any similar pursuit requiring a good deal of skill, transformational calculations must occur before the effective response is achieved. Once the "move" has been made, it may arouse an "aha" reaction: The move was successful, or, "Oh dear, why didn't I see that!" Again the physiological changes that occur are those that have been shown in the Berdina experiments, although the theories of Jennings et al. (1971) and Obrist et al. (1970) would predict the opposite.

Some of the task variables which affect arousal and activation and the effort required to overcome resistance to extinction of a prior set during categorizing and reasoning have also been investigated by Dahl and Spence (1971). They propose an activation

theory which adopts Bergum's (1966) technique of taxonomic analysis of performance, in which tasks are evaluated according to complexity and activation of response. Task demands were rated by independent judges, and the amount of heart rate change was correlated to each task. Heart rate increased with response demand, and all correlations were significant. A subsequent factor analysis revealed a major factor which they designated "density," or degree of concentration required; in other words, the total amount of cognitive *effort* involved in the task. The other factor which accounted for the greatest amount of variance in task performance was information content of relevant stimuli. Thus, again, collative variables are distinguished from the amount of transformational effort or "work" required in problem solution.

Cognitive Effort

Do these results mean that cognitive effort, the effort of thinking and problem solving, is due entirely to the activation of peripheral muscular mechanisms as reflected in changes of heart rate? We cannot answer this question conclusively at this time, and in fact, there may be no conclusive answer, at least not where normal intact problem-solving organisms are concerned. The neuromuscular apparatus works as a unit, brain and body inseparably yoked. Yet Malmö (Malmö & Belanger, 1967) and Lacey (B. C. Lacey & J. I. Lacey, 1974) have made an excellent case in their extensive, detailed, and perceptive reviews that there is good evidence to believe that the effort involved in cognitive work is not wholly due to changes observed in the muscular system. Further evidence comes from pathological situations such as complete cervical spinal transections which arise when partial separation between brain and body occur. We have already reviewed the fact that theta activity in the hippocampal formation, usually present when the animal moves about voluntarily, is also recorded under the same circumstances in curarized and therefore completely paralyzed subjects. A broken neck that completely severs the muscular system from control by the brain does not pre-

clude effortful problem solving. The competency to think remains.

We know from observations of sensory disturbances that processes which are initially localized peripherally become progressively neuralized. Thus phantom limb pain following an amputation can be successfully treated immediately by injections of an anaesthetic at the site of amputation. After a few months, however, such peripheral injections usually fail to have an effect, and spinal cordotomy (section of the spinothalamic tracts) must be done to relieve the disturbance. Should several years of intractable pain be experienced before treatment, even spinal cordotomy often proves useless. The brain itself must be operated upon: Frontal leukotomy or parietal lobectomy are now resorted to.

It is thus plausible that the neuromuscular processes demanding effort are to a large extent initially peripheral in their origin and manifestation. But as problem solving progresses and problem-solving skill develops, the effort becomes more and more a concomitant of the brain processes involved—apart from, or only *reflected in*, the peripheral manifestations—and thus becomes truly cognitive. How then are we to study and understand the brain processes that demand effort?

Studies of local differences in cerebral blood flow and metabolism provide the obvious answer (for review see Ingvar and Lassen, in press). However, it is also helpful to define the issues involved, especially since overall cerebral metabolism is augmented so little by even the most effortful problem-solving activities. Just as quantifiable definitions of information and uncertainty become useful in understanding information processing, so quantitative definitions of effort, work, and energy become useful in understanding control mechanisms. We have already noted that effort can be defined as the measure of attention "paid" to increase or maintain the efficiency of a communication channel by reducing its equivocation. As detailed, such reduction can be accomplished by reducing noise or changing the structure of the redundancy of the communication—chunking or parsing being

prime examples of such a process. McFarland (1971) has approached the problem from a control rather than an information theoretic stance and comes to the identical conclusion. In keeping with engineering tradition, he defines work (pp. 215–220) as the measure of the rate of changing the state of a system and energy as the capacity for doing work. Thus the changes in neuronal representational model, changes in information processing competency, in set, attitude or state, developed by the hippocampal circuit described in this review, demand work—by definition. Further, resistance to change may occur when changes take place in related systems and the changes are not independent of one another: For example, stimulus and response may originally be coupled as in a defense reaction—that is, internally constrained by a preexisting neuronal model. In such a situation, a change in state is produced only when the input and output systems become uncoupled by the hippocampal circuit. In a distraction situation, maintenance of state involves loosening an external constraint as we have seen, and the uncoupling of arousal and activation, thus increasing internal constraint. As an analogy, change in blood sugar level may be dependent upon, or constrained by, change in osmotic pressure and basal temperature and motor activity so that the behavior of an organism must be analyzed in relation to such constraints (see for example, McFarland, 1971).

The constraints operating in a system can often be expressed in terms of integrable differential or algebraic equations, in which case they are known as "holonomic constraints." Holonomically constrained systems allow description in terms of "the number of generalized coordinates required to describe the state of the system and this number is known as the 'degrees of freedom' of the system" (McFarland, 1971, p. 217). In systems showing holonomic constraints, it should therefore be possible, in principle, to measure resistance and therefore the effort necessary to overcome the resistance. For instance, in discrimination reversal and in delayed alternation, the rate of extinction of the hippocampal theta peak to the previously rewarded

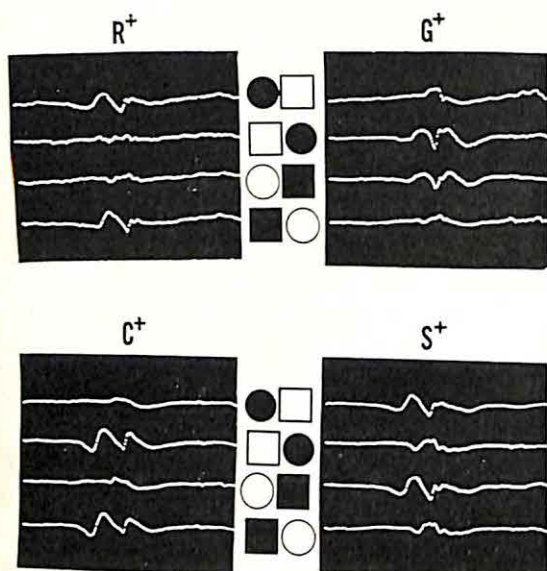


FIGURE 10. Comparison of *response-locked* activity evoked in temporal cortex when monkeys are performing (90% correct) color (top panels) and pattern (bottom panels) discrimination. The average evoked response (300 trials) was produced by the stimulus configuration shown in the diagrams between the panels. Each tracing = 250 msec prior to and 250 msec just after each response. Note that during the color discriminations the 1st and 4th (and the 2nd and 3rd) traces are similar, while during pattern discriminations the 1st and 3rd (and 2nd and 4th) traces are alike. These similarities reflect the position of the color cues in the color task and the position of the patterns in the pattern task. Position per se, however, is not encoded in these traces. Note that this difference occurs despite the fact that the retinal image formed by the flash stimulus is identical in the pattern and color problems. (Abbreviations: R⁺ = Red reinforced; G⁺ = Green reinforced; C⁺ = Circle reinforced; S⁺ = Square reinforced.)

stimulus might be used quantitatively in measuring the resistance involved in switching attention from one cue to another. In a selective attention experiment, where first one then another dimension of a multidimensional cue is reinforced (see Figure 10; Pribram, Day, & Johnston, in press; Rothblatt & Pribram, 1972), recording of hippocampal theta and mapping cortical CNV or desynchronization during the changeover from one dimension to the other might be similarly useful. Neurologically, this measure of resistance would reflect the rate of release from entrainment of rhythmic processes by a temporary dominant focus or

pacemaker. The problem is at the forefront of our capacity to conceptualize it; however, the direction in which conceptualization must proceed has been given: We must ascertain whether the constraints operating in a particular behavioral system are holonomically related and determine the number of coordinates that are required to describe it (the degrees of freedom involved). We can then perform rigorous quantitative experimental analyses of changes of state (transformations) of each of the coordinate subsystems and thus finally arrive at the resistance of the system as a whole and the effort involved in overcoming this resistance. E. Roy John (Bartlett & John, 1973; John, Bartlett, Shimokochi, & Kleinman, 1973), Adey (1967), Callaway and Harris (1974), and our laboratory (Pribram, Day & Glick, in press) have begun this type of analysis and initial results, at least, are promising.

This review has dealt with initial specification of the locus of resistance in the attention mechanism. We have outlined in a very preliminary way that at least two major subsystems, a visceromotor "arousal" and a somatomotor "activation" mechanism, must be considered. We have also reviewed the evidence for a control mechanism operating on each system (the amygdala circuits for arousal; basal ganglia for activation) and a mechanism (the hippocampal circuit) that operates to overcome the constraints which would ordinarily maintain a tight join between stimulus and response and thus keep these two control systems from functioning. We have suggested that overcoming these constraints may demand effort. Whether the constraints in attention operate holonomically or nonholonomically remains to be determined; but the techniques for determination are available. Should it turn out that constraints can be expressed or even approximated in terms of integrable differential equations, a measure on cognitive effort would be attainable.

Whether the attention comes by grace of genius or by dint of will, the longer one does attend to a topic the more mastery of it one has. And the faculty of voluntarily bringing back a wandering attention, over and over again, is the very root of judgment, character, and will (James, 1950, p. 424).

REFERENCES

- Abrahams, V. C., & Hilton, S. M. Active muscle vasodilation and its relation to the "fight and flight reactions" in the conscious animal. *Journal of Physiology*, 1958, 140, 16-17.
- Abrahams, V. C., Hilton, S. M., & Zbrozyna, A. W. The role of active muscle vasodilation in the alerting stage of the defense reaction. *Journal of Physiology*, 1964, 171, 189-202.
- Adey, W. R. Intrinsic organization of cerebral tissue in alerting, orienting, and discriminative responses. In G. C. Quarton, T. Melnechuck, & F. O. Schmitt (Eds.), *The neurosciences: A study program*. New York: Rockefeller University Press, 1967.
- Adey, W. R. Spontaneous electrical brain rhythms accompanying learned responses. In F. O. Schmitt (Ed.), *The neurosciences second study program*. New York: Rockefeller University Press, 1970.
- Adey, W. R., Dunlop, C. W., & Hendrix, C. E. Hippocampal slow-waves: Distribution and phase relations in the course of approach learning. *Archives of Neurology*, 1960, 3, 96-112.
- Anderson, N. S., & Fitts, P. M. Amount of information gained during brief exposures of numerals and colors. *Journal of Experimental Psychology*, 1958, 56, 362-369.
- Bagshaw, M. H., & Benzie, S. Multiple measures of the orienting reaction and their dissociation after amygdectomy in monkeys. *Experimental Neurology*, 1968, 20, 175-187.
- Bagshaw, M. H., & Coppock, H. W. Galvanic skin response conditioning deficit in amygdectomized monkeys. *Experimental Neurology*, 1968, 20, 188-196.
- Bagshaw, M. H., Kimble, D. P., & Pribram, K. H. The GSR of monkeys during orienting and habituation and after ablation of the amygdala, hippocampus and inferotemporal cortex. *Neuropsychologia*, 1965, 3, 111-119.
- Bagshaw, M. H., & Pribram, J. D. Effect of amygdectomy on stimulus threshold of the monkey. *Experimental Neurology*, 1968, 20, 197-202.
- Bartlett, F., & John, E. R. Equipotentiality quantified: The anatomical distribution of the engram. *Science*, 1973, 181, 764-767.
- Berdina, N. A., Kolenko, O. L., Kotz, I. M., Kuzetzov, A. P., Rodinov, I. M., Savtchenko, A. P., & Thorevsky, V. I. Increase in skeletal muscle performance during emotional stress in man. *Circulation Research*, 1972, 6, 642-650.
- Bergum, B. O. A taxonomic analysis of continuous performance. *Perceptual and Motor Skills*, 1966, 23, 47-54.
- Berlyne, D. E. The development of the concept of attention in psychology. In C. R. Evans & T. B. Mulholland (Eds.), *Attention in neurophysiology*. New York: Appleton-Century-Crofts, 1969.
- Black, A. H., & Young, G. A. Electrical activity of the hippocampus and cortex in dogs operantly trained to move and to hold still. *Journal of Comparative and Physiological Psychology*, 1972, 79, 128-141.
- Black, A. H., Young, G. A., & Batenchuck, C. The avoidance training of hippocampal theta waves in flaxedized dogs and its relation to skeletal movement. *Journal of Comparative Physiological Psychology*, 1970, 70, 15-24.
- Bland, B. H., & Vanderwolf, C. H. Diencephalic and hippocampal mechanisms of motor activity in the rat: Effect of posterior hypothalamic stimulation on behavior and hippocampal slow wave activity. *Brain Research*, 1972, 43, 67-88. (a)
- Bland, B. H., & Vanderwolf, C. H. Electrical stimulation of the hippocampal formation: Behavioral and bioelectrical effects. *Brain Research*, 1972, 43, 89-106. (b)
- Blaylock, B. Some antecedents of directional fractionation: Effects of "intake-rejection", verbalization requirements, and threat of shock on heart rate and skin conductance. *Psychophysiology*, 1972, 9, 40-52.
- Bonvallet, M., & Allen, M. B. Prolonged spontaneous and evoked reticular activation following discrete bulbar lesions. *Electroencephalography and Clinical Neurophysiology*, 1963, 15, 969-988.
- Bremner, F. J. The effect of habituation and conditioning trials on hippocampal EEG. *Psychonomic Science*, 1970, 18, 181-183.
- Brillouin, L. *Science and information theory*. New York: Academic Press, 1962.
- Brindley, G. S., & Merton, P. A. The absence of position sense in the human eye. *Journal of Physiology*, 1960, 153, 127-130.
- Broadbent, D. E. *Decision and stress*. New York: Academic Press, 1971.
- Bruner, J. S. On perceptual readiness. *Psychological Review*, 1957, 64, 123-152.
- Bull, N. The attitude theory of emotion. *Nervous and Mental Disease Monographs*, 1951, No. 81.
- Bull, K., & Lang, P. J. Intensity judgments and physiological response amplitudes. *Psychophysiology*, 1972, 9, 428-436.
- Bures, J., & Buresova, O. Plasticity in single neurones and neural populations. In G. Horn & R. A. Hinde (Eds.), *Short-term changes in neural activity and behaviour*. Cambridge, England: Cambridge University Press, 1970.
- Butter, C. M. The effect of discrimination training on pattern equivalence in monkeys with inferotemporal and lateral striate lesions. *Neuropsychologia*, 1968, 6, 27-40.
- Butter, C. M. Impairments in selective attention to visual stimuli in monkeys with inferotemporal and lateral striate lesions. *Brain Research*, 1969, 12, 374-383.
- Callaway, E., & Harris, P. R. Coupling between cortical potentials from different areas. *Science*, 1974, 183, 873-875.
- Chow, K. L. Integrative functions of the thalamocortical visual system of cat. In K. H. Pribram

- & D. Broadbent (Eds.), *Biology of memory*. New York: Academic Press, 1970.
- Clemente, C. C., Serman, M. B., & Wyerwice, W. Post-reinforcement EEG synchronization during alimentary behavior. *Electroencephalography and Clinical Neurophysiology*, 1964, 16, 355-365.
- Crowne, D. P., Konow, A., Drake, K. J., & Pribram, K. H. Hippocampal electrical activity in the monkey during delayed alternation problems. *Electroencephalography and Clinical Neurophysiology*, 1972, 33, 567-577.
- Crowne, D. P., & Riddell, W. I. Hippocampal lesions and the cardiac component of the orienting response in the rat. *Journal of Comparative and Physiological Psychology*, 1969, 69, 748-755.
- Dahl, H., & Spence, D. P. Mean heart rate predicted by task demand characteristics. *Psychophysiology*, 1971, 7, 369-376.
- Dalton, A., & Black, A. H. Hippocampal electrical activity during the operant conditioning of movement and refraining from movement. *Communications in Behavioral Biology*, 1968, 2, 267-273.
- Donchin, E., Gerbrandt, L. A., Leifer, L., & Tucker, L. Is the contingent negative variation contingent on a motor response? *Psychophysiology*, 1972, 9, 178-188.
- Donchin, E., Otto, D., Gerbrandt, L. K., & Pribram, K. H. While a monkey waits: Electrocardiac events recorded during the foreperiod of a reaction time study. *Electroencephalography and Clinical Neurophysiology*, 1971, 31, 115-127.
- Donchin, E., Otto, D., Gerbrandt, L. K., & Pribram, K. H. While a monkey waits. In K. H. Pribram & A. R. Luria (Eds.), *Psychophysiology of the frontal lobes*. New York: Academic Press, 1973.
- Douglas, R. J., & Pribram, K. H. Learning and limbic lesions. *Neuropsychologia*, 1966, 4, 197-220.
- Douglas, R. J., & Pribram, K. H. Distraction and habituation in monkeys with limbic lesions. *Journal of Comparative Physiological Psychology*, 1969, 69, 473-480.
- Eckberg, D. L., Fletcher, G. F., & Braunwald, E. Mechanism of prolongation of the R-R interval with electrical stimulation of the carotid sinus nerves in man. *Circulation Research*, 1972, 30, 131-138.
- Elazar, A., & Adey, W. R. Spectral analysis of low frequency components in the electrical activity of the hippocampus during learning. *Electroencephalography and Clinical Neurophysiology*, 1967, 23, 225-240.
- Elliott, R. Tonic heart rate: Experiments on the effects of collative variables lead to a hypothesis about its motivational significance. *Journal of Personality and Social Psychology*, 1969, 12, 211-228.
- Elliott, R., Bankart, B., & Light, T. Differences in the motivational significance of heart rate and palmar conductance: Two tests of a hypothesis. *Journal of Personality and Social Psychology*, 1970, 14, 166-172.
- Feldman, S. M., & Waller, H. J. Dissociation of electrocortical activation and behavior arousal. *Nature*, 1962, 196, 1320.
- Fibiger, H. C., Phillips, A. G., & Clouston, R. A. Regulatory deficits after unilateral electrolytic or 6-OHDA lesions of the substantia nigra. *American Journal of Physiology*, 1973, 225, 1282-1287.
- Forsyth, R. P. Hypothalamic control of the distribution of cardiac output in the unanesthetized Rhesus monkey. *Circulation Research*, 1970, 26, 783-794.
- Fuller, J. L., Rosvold, H. E., & Pribram, K. H. The effect on affective and cognitive behavior in the dog of lesions of the pyriform-amygdala-hippocampal complex. *Journal of Comparative and Physiological Psychology*, 1957, 50, 89-96.
- Galambos, R., Norton, T. T., & Frommer, C. P. Optic tract lesions sparing pattern vision in cats. *Experimental Neurology*, 1967, 18, 8-25.
- Garner, W. R. *Uncertainty and structure as psychological concepts*. New York: Wiley, 1962.
- Gastaut, H. Étude électrographique chez l'homme et chez l'animal des décharges épileptiques dites "psychométriques". *Revue Neurologique*, 1953, 8, 310-354.
- Gastaut, H. Interpretation of the symptoms of "psychomotor" epilepsy in relation to physiologic data on Rhinencephalic function. *Epilepsia*, 1954 (Series 3), 3, 84-88.
- Germana, J. Response characteristics and the orienting reflex. *Journal of Experimental Psychology*, 1968, 78, 610-616.
- Germana, J. Autonomic-behavioral integration. *Psychophysiology*, 1969, 6, 78-90.
- Gloor, P. Études électrographiques de certaines connexions rhinencephaliques. In Ph. Alajouanine (Ed.), *Physiologie et pathologie du rhinencephale*. Paris: Masson, 1961.
- Graham, F. K., & Clifton, R. K. Heart rate changes as a component of the orienting response. *Psychological Bulletin*, 1966, 65, 305-320.
- Grandstaff, N. W. Frequency analysis of EEG during milk drinking. *Electroencephalography and Clinical Neurophysiology*, 1969, 27, 55-57.
- Grandstaff, N. W., & Pribram, K. H. Habituation: Electrical changes in the visual system. *Neuropsychologia*, 1972, 10, 125-132.
- Grastyan, E. The hippocampus and higher nervous activity. In M. A. B. Brazier (Ed.), *The central nervous system and behavior*. New York: Josiah Macy, Jr. Foundation, 1959.
- Grastyan, E., Lissak, K., Madarasz, I., & Donohoffer, H. Hippocampal electrical activity during the development of conditioned reflexes. *Electroencephalography and Clinical Neurophysiology*, 1959, 11, 409-430.
- Gray, J. *The psychology of fear and stress*. New York: McGraw-Hill, 1971.

- Green, J. F., & Arduini, A. Hippocampal electrical activity in arousal. *Journal of Neurophysiology*, 1954, 17, 533-557.
- Groves, P. M., & Thompson, R. F. Habituation: A dual-purpose theory. *Psychological Review*, 1970, 77, 419-450.
- Grueninger, W. E., & Grueninger, J. The primate frontal cortex and allasostasis. In K. H. Pribram & A. R. Luria (Eds.), *Psychophysiology of the frontal lobes*. New York: Academic Press, 1973.
- Grueninger, W. E., Kimble, D. P., Grueninger, J., & Levine, S. GSR and corticosteroid response in monkeys with frontal ablations. *Neuropsychologia*, 1965, 3, 205-216.
- Gruzelier, J. H., & Venables, P. H. Skin conductance orienting activity in a heterogeneous sample of schizophrenics. *Journal of Nervous and Mental Disease*, 1972, 155, 277-287.
- Hahn, W. W. Attention and heart rate: A critical appraisal of the hypothesis of Lacey and Lacey. *Psychological Bulletin*, 1973, 79, 59-70.
- Haider, M. Neuropsychology of attention, expectation, and vigilance. In D. I. Mostofsky (Ed.), *Attention: Contemporary theory and analysis*. New York: Appleton-Century-Crofts, 1970.
- Hare, R. D. Response requirements and directional fractionation of autonomic responses. *Psychophysiology*, 1972, 9, 419-427.
- Helmholtz, H. v. *Treatise on physiological optics* (Vol. 1). Menasha, Wis.: Optical Society of America, 1924.
- Helson, H. *Adaptation-level theory*. New York: Harper & Row, 1964.
- Hillyard, S. A. The CNV and human behavior: A review. In W. C. McCallum & J. R. Knott (Eds.), *Event related slow potentials of the brain: Their relation to behavior*. Amsterdam: Elsevier, 1973.
- Horn, G. Changes in neuronal activity and their relationship to behaviour. In G. Horn & R. A. Hinde (Eds.), *Short-term changes in neural activity and behaviour*. Cambridge, England: Cambridge University Press, 1970.
- Horn, G., & Hill, R. M. Responsiveness to sensory stimulation of units in the superior colliculus and subjacent tectotegmental regions of the rabbit. *Experimental Neurology*, 1966, 14, 199-224.
- Horn, G., & Hinde, R. A. (Eds.), *Short-term changes in neural activity and behaviour*. Cambridge, England: Cambridge University Press, 1970.
- Ingvar, D. H., & Lassen, N. A. (Eds.), *The working brain: The coupling of function, metabolism and blood flow in the brain*. Copenhagen: Alfred Benzon Foundation, in press.
- James, W. *Principles of psychology* (Vol. 1). New York: Dover, 1950.
- Jennings, J. R., Averill, J. R., Opton, E. M., & Lazarus, R. S. Some parameters of heart rate change: Perceptual versus motor task requirements, noxiousness, and uncertainty. *Psychophysiology*, 1971, 7, 194-212.
- Johansson, G., & Frankenhauser, M. Temporal factors in sympatho-adrenomedullary activity following acute behavioral activation. *Biological Psychology*, 1973, 1, 63-73.
- John, E. R. *Mechanisms of memory*. New York: Academic Press, 1967.
- John, E. R., Bartlett, F., Shimokochi, M., & Kleinman, D. Neural readout from memory. *Journal of Neurophysiology*, 1973, 36, 893-924.
- John, E. R., & Killam, K. F. Electrophysiological correlates of differential approach-avoidance conditioning in cats. *Journal of Nervous and Mental Diseases*, 1960, 131, 183-201.
- Jung, R., & Kornmueller, A. E. Eine methodik der Abkütung lokalisierter Potentialschwankungen aus subcorticalen Hirngebieten. *Arch für Psychiatrie und Nervenkrankheiten*, 1938, 109, 1-30.
- Kaada, B. R., Pribram, K. H., & Epstein, J. A. Respiratory and vascular responses in monkeys from temporal pole, insula, orbital surface and cingulate gyrus: A preliminary report. *Journal of Neurophysiology*, 1949, 12, 347-356.
- Kahneman, D. *Attention and effort*. Englewood Cliffs, N. J.: Prentice-Hall, 1973.
- Kilpatrick, D. G. Differential responsiveness of two electrodermal indices to psychological stress and performance of a complex cognitive task. *Psychophysiology*, 1972, 9, 218-226.
- Kimble, D. P., Bagshaw, M. H., & Pribram, K. H. The GSR of monkeys during orienting and habituation after selective partial ablations of the cingulate and frontal cortex. *Neuropsychologia*, 1965, 3, 121-128.
- Koepke, J. E., & Pribram, K. H. Habituation of the GSR as a function of stimulus duration and "spontaneous activity". *Journal of Comparative and Physiological Psychology*, 1966, 3, 442-448.
- Koepke, J. E., & Pribram, K. H. Habituation of the vasoconstriction response as a function of stimulus duration and anxiety. *Journal of Comparative and Physiological Psychology*, 1967, 64, 502-504.
- Konrad, K. W., & Bagshaw, M. H. Effect of novel stimuli on cats reared in a restricted environment. *Journal of Comparative and Physiological Psychology*, 1970, 70, 157-164.
- Kornhuber, H. H., & Deecke, L. Hirnpotentialänderungen bei Willkurbewegungen und passiven Bewegungen des Menschen: Bereitschaftspotential und reafferente Potentiale. *Pflügers Archiv für die gesamte Physiologie des Menschen und der Tiere*, 1965, 284, 1-17.
- Lacey, B. C., & Lacey, J. I. Studies of heart rate and other bodily processes in sensorimotor behavior. In P. A. Obrist, A. Black, J. Bruner, & L. DiCara (Eds.), *Cardiovascular psychophysiology: Current issues in response mechanisms, biofeedback and methodology*. Chicago: Aldine-Atherton, 1974.
- Lacey, J. I. Somatic response patterning and stress: Some revisions of activation theory.

- In M. H. Appley & R. Trumball (Eds.), *Psychological stress: Issues in research*. New York: Appleton-Century-Crofts, 1967.
- Lacey, J. I., & Lacey, B. C. Some autonomic central nervous system interrelationships. In P. Black (Ed.), *Physiological correlates of emotion*. New York: Academic Press, 1970.
- Lacey, J. I., & Lacey, B. C. On heart rate responses and behavior: A reply to Elliott. *Journal of Personality and Social Psychology*, 1974, 30, 1-18.
- Lader, M. H. The effects of cyclobarbitone on the pulse volume, pulse rate and electromyogram. *Journal of Psychosomatic Research*, 1965, 8, 385-398.
- Lashley, K. S. *Brain mechanisms and intelligence*. Chicago: University of Chicago Press, 1929.
- Lawrence, D. H. Acquired distinctiveness of cues: I. Transfer between discriminations on the basis of familiarity with the stimulus. *Journal of Experimental Psychology*, 1949, 39, 776-784.
- Lawrence, D. H. Acquired distinctiveness of cues: II. Selective association in a constant stimulus situation. *Journal of Experimental Psychology*, 1950, 40, 175-188.
- Lewis, M., & Wilson, C. D. The cardiac response to a perceptual-cognitive task in the young child. *Psychophysiology*, 1970, 6, 411-420.
- Lindsay, P. H. Multichannel processing in perception. In D. I. Mostofsky (Ed.), *Attention: Contemporary theory and analysis*. New York: Appleton-Century-Crofts, 1970.
- Lindsley, D. B. The reticular activating system and perceptual integration. In D. E. Sheer (Ed.), *Electrical stimulation of the brain*. Austin: University of Texas Press, 1961.
- Luria, A. R., & Homskaya, E. D. Frontal lobes and the regulation of arousal processes. In D. I. Mostofsky (Ed.), *Attention: Contemporary theory and analysis*. New York: Appleton-Century-Crofts, 1970.
- Luria, A. R., Pribram, K. H., & Homskaya, E. D. An experimental analysis of the behavioral disturbance produced by a left frontal arachnoidal endothelloma (meningioma). *Neuropsychologia*, 1964, 2, 257-280.
- Macadar, A. W., Chalupa, L. M., & Lindsley, D. B. Differentiation of brain stem loci which affect hippocampal and neocortical electrical activity. *Experimental Neurology*, 1974, 43, 499-514.
- Mackworth, J. F. *Vigilance and habituation*. Harmondsworth, Middlesex, England: Penguin, 1969.
- MacLean, P. D. Psychosomatic disease and the "visceral brain": Recent developments bearing on the Papez theory of emotion. *Psychosomatic Medicine*, 1949, 11, 338-353.
- Magoun, H. W. *The waking brain*. Springfield, Ill.: Charles C Thomas, 1958.
- Malmo, R. B., & Bélanger, D. Related physiological and behavioral changes: What are their determinants? In H. C. Solomon, S. Cobb, & W. Penfield (Eds.), *The brain and human behavior*. Baltimore, Md.: Williams & Wilkins, 1967.
- Mandler, G. The interruption of behavior. In D. Levine (Ed.), *Nebraska Symposium on Motivation* (Vol. 12). Lincoln: University of Nebraska Press, 1964.
- Matthews, A. M., & Lader, M. H. An evaluation of forearm blood flow as a psychophysiological measure. *Psychophysiology*, 1971, 8, 509-524.
- McFarland, D. J. *Feedback mechanisms in animal behaviour*. New York: Academic Press, 1971.
- McGuinness, D. Cardiovascular responses during habituation and mental activity in anxious men and women. *Biological Psychology*, 1973, 1, 115-123.
- McNeil, M. *Studies in the biology of emotion and motivation*. Unpublished doctoral dissertation, Stanford University, 1972.
- Mednick, S. A., & Schulsinger, F. Some pre-morbid characteristics related to breakdown in children with schizophrenic mothers. In D. Rosenthal & S. S. Kety (Eds.), *The transmission of schizophrenia*. New York: Pergamon Press, 1968.
- Merton, P. A. Speculations on the servo-control of movement. In J. L. Malcolm & J. A. B. Gray (Eds.), *The spinal cord*. London: J. & A. Churchill, Ltd., 1953.
- Mettler, A., & McLardy, T. Posterior cuts in prefrontal leucotomy: A clinicopathological study. *Journal of Mental Science*, 1948, 94, 555-564.
- Mettler, F. A., & Mettler, C. C. The effects of striatal injury. *Brain*, 1942, 65, 242-255.
- Miller, G. A. The magical number seven, plus or minus two, or, some limits on our capacity for processing information. *Psychological Review*, 1956, 63, 81-97.
- Miller, G. A., Galanter, E. H., & Pribram, K. H. *Plans and the structure of behavior*. New York: Holt, Rinehart & Winston, 1960.
- Milner, B. Psychological defects produced by temporal lobe excision. In H. C. Solomon, S. Cobb, & W. Penfield (Eds.), *The brain and human behavior*. Baltimore, Md.: Williams & Wilkins, 1958.
- Mishkin, M., & Pribram, K. H. Analysis of the effects of frontal lesions in monkeys: I. Variations of delayed alternation. *Journal of Comparative Physiological Psychology*, 1955, 48, 492-495.
- Mostofsky, D. I. (Ed.), *Attention: Contemporary theory and analysis*. New York: Appleton-Century-Crofts, 1970.
- Neisser, U. *Cognitive psychology*. New York: Appleton-Century-Crofts, 1967.
- Obrist, P. A., Howard, J. L., Hennis, H. S., & Murrell, D. J. Cardiac-somatic changes during simple reaction time tasks: A developmental study. *Psychophysiology*, in press.
- Obrist, P. A., Howard, J. L., Lawler, J. E., Sutterer, J. R., Smithson, W. W., & Martin, P. L. Alternations in cardiac contractibility during

- classical aversive conditioning in dogs: Methodological and theoretical implications. *Psychophysiology*, 1972, 9, 246-261.
- Obrist, P. A., Webb, R. A., Sutterer, J. R., & Howard, J. L. Cardiac deceleration and reaction time: An evaluation of two hypotheses. *Psychophysiology*, 1970, 6, 695-706. (a)
- Obrist, P. A., Webb, R. A., Sutterer, J. R., & Howard, J. L. The cardiac-somatic relationship: Some reformulations. *Psychophysiology*, 1970, 6, 569-587 (b)
- Obrist, P. A., Wood, D. M., & Perez-Reyes, M. Heart rate during conditioning in humans: Affect of UCS intensity vagal blockade and adrenergic block of vasomotor activity. *Journal of Experimental Psychology*, 1965, 70, 32-42.
- Petro, J. K., Holland, A. P., & Bouman, L. N. Instantaneous cardiac acceleration in man induced by a voluntary muscle contraction. *Journal of Applied Physiology*, 1970, 24, 794-798.
- Picton, T. W., & Hillyard, S. A. Human auditory evoked potentials. II. Effects of attention. *Electroencephalography and Clinical Neurophysiology*, 1974, 36, 191-200.
- Pribram, K. H. The intrinsic systems of the forebrain. In J. Field, H. W. Magoun, & V. E. Hall (Eds.), *Handbook of physiology, neurophysiology II*. Washington, D. C.: American Physiological Society, 1960. (a)
- Pribram, K. H. A review of theory in physiological psychology. In *Annual review of psychology* (Vol. 11) Palo Alto, Calif. Annual Reviews, 1960. (b)
- Pribram, K. H. Limbic system. In D. E. Sheer (Ed.), *Electrical stimulation of the brain*. Austin: University of Texas Press, 1961.
- Pribram, K. H. Interrelations of psychology and the neurological disciplines. In S. Koch (Ed.), *Psychology: A study of a science* (Vol. 4). *Biologically oriented fields: Their place in psychology and in biological sciences*. New York: McGraw-Hill, 1962.
- Pribram, K. H. Reinforcement revisited: A structural view. In M. Jones (Ed.), *Nebraska Symposium on Motivation* (Vol. 11). Lincoln: University of Nebraska Press, 1963.
- Pribram, K. H. Memory and the organization of attention. In D. B. Lindsley & A. A. Lumsdaine (Eds.), *Brain function* (Vol. 4). Berkeley: University of California Press, 1967.
- Pribram, K. H. Neural servosystems and the structure of personality. *Journal of Nervous and Mental Disease*, 1969, 140, 30-39.
- Pribram, K. H. *Languages of the brain: Experimental paradoxes and principles in neuropsychology*. Englewood Cliffs, N. J.: Prentice-Hall, 1971.
- Pribram, K. H., & Bagshaw, M. H. Further analysis of the temporal lobe syndrome utilizing frontotemporal ablations in monkeys. *Journal of Comparative Neurology*, 1953, 99, 347-375.
- Pribram, K. H., & Broadbent, D. E. (Eds.), *Biology of memory*. New York: Academic Press, 1970.
- Pribram, K. H., Day, R., & Glick, J. Some current methods in neuropsychological research. In S. Bogoch (Ed.), *The future of the brain sciences* (Vol. 2). New York: Plenum Press, in press.
- Pribram, K. H., Day, R., & Johnston, V. S. Selective attention: Distinctive brain electrical patterns produced by differential reinforcement in monkey and man. In D. I. Mostofsky (Ed.), *Behavior control and modification of physiological activity*, in press.
- Pribram, K. H., Douglas, R. J., & Pribram, B. J. The nature of nonlimbic learning. *Journal of Comparative and Physiological Psychology*, 1969, 69, 765-772.
- Pribram, K. H., & Kruger, L. Functions of the "olfactory brain". *Annals of the New York Academy of Sciences*, 1954, 58, 109-138.
- Pribram, K. H., Lim, H., Poppen, R., & Bagshaw, M. H. Limbic lesions and the temporal structure of redundancy. *Journal of Comparative and Physiological Psychology*, 1966, 61, 368-373.
- Pribram, K. H., Reitz, S., McNeil, M., & Spevack, A. A. The effect of amygdectomy on orienting and classical conditioning. In *Mechanisms of formation and inhibition of conditional reflex* (Asratyan Festschrift). Moscow: Publications Office, Academy of Sciences of USSR "Nauka," 1974.
- Pribram, K. H., Spinelli, D. N., & Kamback, M. C. Electrocorical correlates of stimulus response and reinforcement. *Science*, 1967, 157, 94-96.
- Pribram, K. H., & Tubbs, W. E. Short-term memory, parsing, and the primate frontal cortex. *Science*, 1967, 156, 1765-1767.
- Radulovacki, M., & Adey, W. R. The hippocampus and the orienting reflex. *Experimental Neurology*, 1965, 12, 68-83.
- Raphelson, A. C., Isaacson, R. L., & Douglas, R. J. The effect of distracting stimuli on the runway performance of limbic damaged rats. *Psychonomic Science*, 1965, 3, 483-484.
- Rebert, C. S. Cortical and subcortical slow potentials in the monkey's brain during a preparatory interval. *Electroencephalography and Clinical Neurophysiology*, 1972, 33, 389-402.
- Rebert, C. S. Elements of a general cerebral system related to CNV genesis. In W. C. McCallum & J. R. Knott (Eds.), *Event related slow potentials of the brain: Their relationships to behavior*. Amsterdam: Elsevier, 1973.
- Rebert, C. S. Slow potential changes in the monkey's brain during reaction time foreperiod. In C. W. McCallum & J. R. Knott (Eds.), *Event related slow potentials of the brain: Their relationships to behavior* (Vol. 2). Amsterdam: Elsevier, in press.
- Riddell, W. L., Rothblat, L. A., & Wilson, W. A., Jr. Auditory and visual distraction in hippo-

- campectomized rats. *Journal of Comparative and Physiological Psychology*, 1969, 67, 216-219.
- Rosell, S., & Uvnäs, B. Vasomotor activity and oxygen uptake in skeletal muscle of the anesthetized cat. *Acta Physiologica Scandinavica*, 1962, 54, 209-222.
- Rosvold, H. E., Mirsky, A. F., & Pribram, K. H. Influence of amygdectomy on social interaction in a monkey group. *Journal of Comparative and Physiological Psychology*, 1954, 47, 173-178.
- Rothblat, L. & Pribram, K. H. Selective attention: Input filter or response selection? *Brain Research*, 1972, 39, 427-436.
- Ruch, T. C., & Shenkin, H. A. The relation of area 13 on orbital surface of frontal lobes to hyperactivity and hyperphagia in monkeys. *Journal of Neurophysiology*, 1943, 6, 349-360.
- Russell, R. W., Singer, G., Flanagan, F., Stone, M., Russell, J. W. Quantitative relations in amygdala modulation of drinking. *Physiology and Behavior*, 1968, 3, 871-875.
- Sauerland, E. K., & Clemente, C. D. The role of the brain stem in orbital cortex induced inhibition of somatic reflexes. In K. H. Pribram & A. R. Luria (Eds.), *Psychophysiology of the frontal lobes*. New York: Academic Press, 1973.
- Schwartzbaum, J. S., Wilson, W. A., Jr., & Morrisette, J. R. The effects of amygdectomy on locomotor activity in monkeys. *Journal of Comparative and Physiological Psychology*, 1961, 54, 334-336.
- Segundo, J. P., & Bell, C. C. Habituation of single nerve cells in the vertebrate nervous system. In G. Horn & R. A. Hinde (Eds.), *Short-term changes in neural activity and behaviour*. Cambridge, England: Cambridge University Press, 1970.
- Sharpless, S., & Jasper, H. Habituation of the arousal reaction. *Brain*, 1956, 79, 655-680.
- Sherrington, C. *Man on his nature*. Garden City, N. Y.: Doubleday, 1955.
- Simon, H. A. How big is a chunk? *Science*, 1974, 183, 482-488.
- Skinner, J. E., & Lindsley, D. B. The nonspecific medio-thalamic-fronto-cortical system: Its influence on electrocortical activity and behavior. In K. H. Pribram & A. R. Luria (Eds.), *Psychophysiology of the frontal lobes*. New York: Academic Press, 1973.
- Sokolov, E. N. Neuronal models and the orienting reflex. In M. A. B. Brazier (Ed.), *The central nervous system and behavior*. New York: Josiah Macy, Jr. Foundation, 1960.
- Sokolov, E. N. *Perception and the conditioned reflex*. New York: MacMillan, 1963.
- Spencer, W. A., & April, R. A. Plastic properties of monosynaptic pathways in mammals. In G. Horn & R. A. Hinde (Eds.), *Short-term changes in neural activity and behavior*. Cambridge, England: Cambridge University Press, 1970.
- Spevack, A., & Pribram, K. H. A decisional analysis of the effects of limbic lesions in monkeys. *Journal of Comparative and Physiological Psychology*, 1973, 82, 211-226.
- Spinelli, D. N., & Pribram, K. H. Changes in visual recovery functions produced by temporal lobe stimulation in monkeys. *Electroencephalography and Clinical Neurophysiology*, 1966, 20, 44-49.
- Spinelli, D. N., & Pribram, K. H. Changes in visual recovery function and unit activity produced by frontal and temporal cortex stimulation. *Electroencephalography and Clinical Neurophysiology*, 1967, 22, 143-149.
- Tecce, J. J. Contingent negative variation (CNV) and psychological processes in man. *Psychological Bulletin*, 1972, 77, 73-108.
- Teitelbaum, P., & Epstein, A. N. The lateral hypothalamus syndrome: Recovery of feeding and drinking after lateral hypothalamic lesions. *Psychological Review*, 1962, 69, 74-90.
- Ungerstedt, U. Brain dopamine neurons and behavior. In F. O. Schmitt & F. G. Worden (Eds.), *The neurosciences third study program*. Cambridge, Mass.: MIT Press, 1974.
- Ursin, H., & Kaada, B. R. Functional localization within the amygdala complex within the cat. *Electroencephalography and Clinical Neurophysiology*, 1960, 12, 120.
- Vanderwolf, C. H. Hippocampal electrical activity and voluntary movement in the rat. *Electroencephalography and Clinical Neurophysiology*, 1969, 26, 407-418.
- Vanderwolf, C. H. Limbic-diencephalic mechanisms of voluntary movement. *Psychological Review*, 1971, 78, 83-113.
- Vaughan, H. G., Costa, L. D., & Ritter, W. Topography of the human motor potential. *Electroencephalography and Clinical Neurophysiology*, 1968, 25, 1-10.
- Venables, P. H., & Christie, M. J. Mechanisms, instrumentation, recording techniques and quantification of response. In W. F. Prokasy & D. C. Raskin (Eds.), *Electrodermal activity in psychological research*. New York: Academic Press, 1973.
- Vinogradova, O. Registration of information and the limbic system. In G. Horn & R. A. Hinde (Eds.), *Short-term changes in neural activity and behavior*. Cambridge, England: Cambridge University Press, 1970.
- Wall, P. D., & Davis, G. D. Three cerebral cortical systems affecting autonomic function. *Journal of Neurophysiology*, 1951, 14, 507-517.
- Walter, W. G. Electrical signs of association, expectancy and decision in the human brain. *Electroencephalography and Clinical Neurophysiology*, 1967, 25, 258-263.
- Walter, W. G., Cooper, R., Aldridge, V. J., McCallum, W. C. & Winter A. L. Contingent negative variation: An electric sign of sensorimotor association and expectancy in the human brain. *Nature*, 1964, 23, 380-384.

- Weinberg, H. The contingent negative variation: Its relation to feedback and expectant attention. *Neuropsychologia*, 1972, 10, 299-306.
- Whishaw, I. Q., Bland, B. H., & Vanderwolf, C. H. Hippocampal activity, behavior, self-stimulation, and heart rate during electrical stimulation of the lateral hypothalamus. *Journal of Comparative and Physiological Psychology*, 1972, 79, 115-127.
- Wicklegren, W. O., & Isaacson, R. L. Effect of the introduction of an irrelevant stimulus on runway performance of the hippocampectomized rat. *Nature*, 1963, 200, 48-50.
- Wilcott, R. C. Skeletal and autonomic inhibition from low-frequency electrical stimulation of the cat's brain. *Neuropsychologia*, in press.
- Wilson, A., & Wilson, A. S. Psychophysiological and learning correlates of anxiety induced muscle relaxation. *Psychophysiology*, 1970, 6, 740-748.
- Wilson, M. Inferotemporal cortex and the processing of visual information in monkeys. *Neuropsychologia*, 1968, 6, 135-140.
- Wood, R. A., & Obrist, P. A. Effects of controlled and uncontrolled respiration on the conditioned heart rate response in humans. *Journal of Experimental Psychology*, 1964, 68, 221.

(Received August 3, 1973; revision
received September 9, 1974)

THEORETICAL NOTES

Generation-Recognition Theory and the Encoding Specificity Principle

Edwin Martin
University of Michigan

Tulving and Thomson recently contradicted the generation-recognition theory of recall with the phrase "recognition failure of recallable words." This phrase, however, is not an admissible summary of their experiments. The word LIGHT in the cue-target pair *head* LIGHT is not the same as the word LIGHT in the free-association pair *dark* LIGHT. Subjects who recall LIGHT to *head* but refuse to identify the LIGHT that goes with *dark* as the same as the LIGHT that goes with *head* are not failing to recognize a recallable word because the word in question is in fact two different words. Tulving and Thomson's argument against generation-recognition theory involves a problematic shift in level of discourse.

According to generation-recognition theory, what recall consists of comes in two parts: (a) The subject scans his memory and comes up with a set of response candidates; and (b) the response candidates in this set are tested against a recognition criterion of some sort to the end that one, some, or all of them are accepted and hence overtly produced. The general class of generation-recognition models is thus concerned with specifying the structures and processes that result in retrieval of a set of plausible responses and with specifying the nature of the screening process that partitions this set into acceptable and unacceptable subsets.

An important feature of this theory is that generation and recognition are ordered processes. The decision criteria that together constitute the recognition test are applied after the retrieval process has supplied a list of response candidates. The importance of this feature lies in the fact that all one needs to do to produce a thoroughly damaging counterexample is to demonstrate recognition failure of recallable words. Indeed, the staged ordering of generation and recognition seems intuitively to be the feature of this theory most directly approachable by experimentation and hence the feature most vulnerable to contradiction. Given this much, plus the fact that a demonstrated

counterexample is always an exciting event in any science, we are seemingly left with the conclusion that Tulving and Thomson (1973) have presented us with an exciting event: They have reported data from three experiments, and on the basis of these they have concluded that nonrecognizable words are recallable and therefore that generation-recognition theory is denied.

The primary purpose of this article is to argue that Tulving and Thomson have not demonstrated recognition failure of recallable words. A logically related secondary purpose is to point out that Tulving and Thomson have blurred two levels of analysis. What I mean by this is that their arguments and conclusions involve undetected or unacknowledged switching between statements about data and statements about theoretical entities, a result of which is a faulty attack on generation-recognition theory.

THE EVIDENCE

The evidence presented by Tulving and Thomson may be schematized as follows: First, the word pair *head* LIGHT was presented to the subject. The subject knew he was later to be tested for his ability to produce LIGHT in response to the cue *head* and subsequently did so with about 60% success. After presentation of the pair *head* LIGHT and before the test for recall of LIGHT to *head*, the same subject was asked to free associate to the stimulus word *dark*. The word LIGHT was among the associates 60-80% of the time. When then asked to recognize the word LIGHT (which he had just generated as an associate of *dark*) as the

Preparation of this article was supported by the Advanced Research Projects Agency of the U. S. Department of Defense and was monitored by the U. S. Air Force Office of Scientific Research under Contract F44620-72-C-0019.

Requests for reprints should be sent to Edwin Martin, Department of Psychology, University of Kansas, Lawrence, Kansas 66045.

TABLE 1
MATERIALS USED IN THE TULVING-THOMSON EXPERIMENTS

List A			List B		
Recall cue	Free-association cue	Target word	Recall cue	Free-association cue	Target word
ground	hot	COLD ^a	hope	low	HIGH ^a
head	dark	LIGHT ^a	stem	long	SHORT ^a
bath	want	NEED ^b	whiskey	lake	WATER
cheese	grass	GREEN	moth	eat	FOOD
stomach	small	LARGE ^a	cabbage	square	ROUND ^a
sun	night	DAY	glass	soft	HARD ^a
pretty	sky	BLUE	country	closed	OPEN ^a
cave	dry	WET ^a	tool	finger	HAND
whistle	tennis	BALL	memory	fast	SLOW ^a
noise	blow	WIND	covering	lining	COAT
glue	table	CHAIR	barn	clean	DIRTY ^a
command	woman	MAN	spider	eagle	BIRD
fruit	bloom	FLOWER	crust	bake	CAKE
home	bitter	SWEET ^a	deep	bed	SLEEP
grasp	infant	BABY ^b	train	white	BLACK ^a
butter	rough	SMOOTH ^a	mountain	leaf	TREE
drink	tobacco	SMOKE	cottage	hate	LOVE
beat	ache	PAIN ^b	art	boy	GIRL
cloth	lamb	SHEEP ^b	adult	labor	WORK ^b
swift	stop	GO	brave	strong	WEAK
lady	king	QUEEN	door	color	RED
blade	scissors	CUT	roll	carpet	RUG ^b
plant	insect	BUG ^b	think	dumb	STUPID ^b
wish	soap	WASH	exist	human	BEING

Note. Adapted from "Encoding Specificity and Retrieval Processes in Episodic Memory" by E. Tulving and D. M. Thomson. *Psychological Review*, 1973, 80, 352-373. Copyright 1973 by the American Psychological Association. Reprinted by permission.

^a Combination of target word and free-association cue names an abstract dimension; combination of target word and recall cue is a specific denotative noun phrase.

^b Combination of target word and free-association cue names a general concept or disposition; combination of target word and recall cue names a concrete manifestation.

original partner of the recall cue *head*, the subject was only 20-30% successful. It is important to note that this 20-30% recognition rate is considerably lower than the approximate 60% recall rate mentioned above. This is important because it is the evidential basis for Tulving and Thomson's conclusion that non-recognized words nevertheless can be recalled.

Tulving and Thomson used 48 words, of which *LIGHT* was only one, and of course the recognition and recall rates varied from word to word. My choice of *LIGHT* was purely expository. But before I defend the proposition that the *LIGHT* example is representative of the whole set of 48 words, let me introduce my general conclusion by again using *LIGHT*. Very simply, the results schematized in the preceding paragraph may be summarized by saying that Tulving and Thomson's subjects were pleased to pair *LIGHT* with *head* and *LIGHT* with *dark*, but reluctant to agree that the *LIGHT* that went with *dark* is the same as the *LIGHT* that went with *head*. In other words, Tulving and

Thomson's subjects sensibly distinguished *LIGHT* (lamp) from *LIGHT* (luminance). True, neither *LIGHT* nor any of the other 47 words they used is as transparently homographic as, say, *PITCHER* (container) - *PITCHER* (baseball), but then subjects are considerably more subtle and flexible than researchers. If this were not true, the memory area would have become theoretically stable a long time ago.

The set of materials used by Tulving and Thomson is reproduced in Table 1. The three columns for each subset give corresponding recall cues (like *head*), free-association cues (like *dark*), and target words (like *LIGHT*). To many of the target words I have added a superscript. Consider first those that have an *a* superscript. When each such target word is combined with its free-association cue, the combination names an abstract dimension of general experience. In contrast, when combined with its recall cue, the combination forms a specific denotative noun phrase. Thus, for example, *white* *BLACK* names a brightness con-

tinuum while *train* BLACK denotes a concrete, imageable thing. Much the same can be said about the target words with a *b* superscript. All of these, in conjunction with their free-association cues, name general concepts or dispositions. When combined with their recall cues, however, they specify concrete manifestations of those concepts or dispositions. Thus *ache* PAIN is a general disposition while *beat* PAIN denotes a specific agentive cause and its effect. The rest of the list is not so easily partitioned in categorical terms, but taken one by one, the same type of analysis can be made. For example, *tobacco* SMOKE must certainly entail a different representational encoding than *drink* SMOKE, the former suggesting something concrete like a cigarette or a pipe and the latter something abstract like a style of life.

The point of all this is that all of the target words in the list are on a par with the LIGHT (lamp)–LIGHT (luminance) case. Accordingly, Tulving and Thomson's summary phrase "recognition failure of recallable words" is inaccurate. It is inaccurate because it states that the words that were not recognized are the same as the words that were recalled, when in fact they are not the same words at all.

The target words in Table 1 are listed from top to bottom in increasing order of success in the recognition test (Tulving & Thomson, 1973, p. 363). Thus the BEING generated as a free associate to *human* was more likely to be recognized as the BEING that went with the recall cue *exist* than was the HIGH generated as a free associate to *low* to be recognized as the HIGH that went with the recall cue *hope*. What this means is that the recall and free-association cues varied in the degree to which they primed similar semantic encodings of their target words. This result highlights a generally unappreciated fact that homography is not a yes-no classification scheme, a point that recently has clearly been made by Anderson and his associates (Anderson & Bower, 1974; Reder, Anderson, & Bjork, 1974). Given that all words are ambiguous to some degree in the absence of a determining context, it follows that encoding specificity effects should be found just about everywhere (e.g., Tulving, 1974).

ENCODING SPECIFICITY

The encoding specificity principle may be stated as follows:

What is stored is determined by what is perceived and how it is encoded, and what is stored determines what retrieval cues are effective in providing

access to what is stored (Tulving & Thomson, 1973, p. 353).

The bulk of the earlier evidence for this principle (see Tulving and Thomson, 1973, for a summary) has been of the sort that if, at the time of input, the target word LIGHT has with it a normatively weak cue like *head*, then in a subsequent cued recall test *head* is a more effective prompt than is a normatively strong substitute cue like *dark*. Whether or not the normative strengths of association of *head* and *dark* with LIGHT have anything to do with cue effectiveness in this situation is another question. The point here is that at the time of input the subject is seen as storing LIGHT (lamp), hence he should certainly be better off in a prompted recall test if LIGHT (lamp) rather than LIGHT (luminance) is what gets primed. And clearly this is what the above quote from Tulving and Thomson says.

Note that the encoding specificity principle talks about encodings—what is stored—not about nominal stimuli and responses. In other words, note that the encoding specificity principle talks about LIGHT (lamp), LIGHT (luminance), PITCHER (container), and PITCHER (baseball), not about LIGHT and PITCHER. Indeed, this distinction is fundamental to the whole idea of encoding specificity. Having noted this, it is obvious that the results of the three experiments reported by Tulving and Thomson are in complete agreement: Subjects are expected to retrieve LIGHT (lamp) to the cue *head* but to resist identifying LIGHT (luminance) as LIGHT (lamp).

What sense, then, are we to make of the summary phrase "recognition failure of recallable words"? Does it mean that LIGHT is simultaneously recallable and not recognizable? Relative to the encoding specificity principle, the answer is "no." This is because the encoding specificity principle does not talk about LIGHT, only about LIGHT (lamp) and LIGHT (luminance). Does it mean that LIGHT (lamp) is simultaneously retrievable and not recognizable? Again the answer must be "no," this time because the experiments did not address the question. In fact, the experiments were designed to guarantee LIGHT (lamp) in retrieval but LIGHT (luminance) in free association and recognition. In short, the summary phrase "recognition failure of recallable words" is antithetical both to encoding specificity as a conceptual notion and to the experimental procedures whose results are justifiably taken to demonstrate encoding specificity.

THE ARGUMENT

Tulving and Thomson's attack on the generation-recognition theory of recall exemplifies a phenomenon worth noting—namely, a shift in level of discourse. Before describing this shift, though, let me restate the feature of generation-recognition theory under attack. A retrieval process is presumed to generate a set of candidates, each of which is potentially translatable into an overt act known as recall. This set of candidates is then evaluated by a recognition process that sorts out the unacceptable candidates. Thus, first comes a retrieval process and then comes a recognition process. If LIGHT (lamp), say, passes through both of these staged processes, then the subject will say LIGHT. If the recognition criterion is set for LIGHT (lamp) and against LIGHT (luminance), the subject will say "no" if presented with a recognition test of LIGHT (luminance).

Now, what about Tulving and Thomson's charge that nonrecognition of recallable words refutes generation-recognition theory? If by "words" they mean entities like LIGHT (lamp), then their experiments, by design, have not addressed the issue, as I pointed out earlier. If by "words" they mean entities like LIGHT, then what they have done is dismissed a critical distinction, a distinction of central importance to encoding specificity theory. This is a shift in level of discourse, a generalization from LIGHT (lamp) and LIGHT (luminance) to just plain LIGHT. This is a curious shift in that both the encoding specificity principle and the experimental results they report argue that this type of generalization is wrong. Moreover, Tulving and Thomson have, by omission, improperly imputed this shift to generation-recognition theorists as well. Kintsch (1970) writes:

The meaning of a formal unit . . . is defined by its relationship with other formal units. That is, a word does not have meaning in itself, but meaning in the system is given entirely by context. . . . The meaning of a word is known only if it is contrasted with the meaning of a related or opposed term (p. 354).

Clearly, the unit of analysis here is LIGHT (lamp), not LIGHT. Thus what Tulving and Thomson have done, contrary both to their own

position and to that of generation-recognition theorists, is move the whole conversation to a level that seems to entail dispute but in fact is vacuous.

Consider the following three statements: (a) LIGHT (lamp) is retrieved and recognized, and the subject said LIGHT. (This is what generation-recognition theory would say about a recallable word.) (b) LIGHT (lamp) was retrieved, and LIGHT (luminance) was not identified as LIGHT (lamp). (This statement is compatible both with the encoding specificity principle and with generation-recognition theory.) (c) LIGHT was recalled but not recognized. (This is the false summary phrase.) Tulving and Thomson experimentally substantiate b but use c in arguing against a.

Undetected or unacknowledged shifts in level of discourse often give rise to irrelevant comparisons and hence pseudodisagreements. I say "pseudo" because they are disagreements that in principle cannot be decided by experimentation, no matter how ingenious. Resolution of such disagreements can come only by referential analysis of the linguistic statements made by the contenders. In the present case, the conclusion is that Tulving's encoding specificity principle, while in itself a cogent and properly influential concept, is currently either unrelated or complementary to the equally sensible view that recall is a staged combination of generation and recognition of response candidates.

REFERENCES

- Anderson, J. R., & Bower, G. H. A propositional theory of recognition memory. *Memory & Cognition*, 1974, 2, 406-412.
- Kintsch, W. Models for free recall and recognition. In D. A. Norman (Ed.), *Models of human memory*. New York: Academic Press, 1970.
- Reder, L. M., Anderson, J. R., & Bjork, R. A. A semantic interpretation of encoding specificity. *Journal of Experimental Psychology*, 1974, 102, 648-656.
- Tulving, E. Recall and recognition of semantically encoded words. *Journal of Experimental Psychology*, 1974, 102, 778-787.
- Tulving, E., & Thomson, D. M. Encoding specificity and retrieval processes in episodic memory. *Psychological Review*, 1973, 80, 352-373.

(Received May 2, 1974)

Features and Semantic Memory: Set-Theoretic or Network Model?

James D. Hollan
Clarkson College of Technology

This brief note attempts to clarify an issue recently raised by Smith, Shoben, and Rips concerning network versus set-theoretic models of semantic memory. It is argued that their set-theoretic model can, without loss of explicatory power, be formulated as a network model and that therefore the distinction is vacuous. Some reasons are also given for preferring network formulations.

Smith, Shoben, and Rips (1974) have proposed a feature model of semantic memory. They contend that the model, in terms of the underlying structural representation, is a set-theoretic rather than a network model, following a distinction formulated earlier (Rips, Shoben, & Smith, 1973). I would like to take this distinction to task. My thesis is that their set-theoretic model can, without loss of explicatory power, be formulated as a network model and that therefore the distinction is vacuous.

Smith et al. (1974) state that "network models assume that words, or their counterparts, exist as independent units in semantic memory connected in a network by labeled relations" (p. 214; Rips et al., 1973, p. 2), while set-theoretic models represent concepts by sets of elements. In the Smith et al. (1974) feature model, the elements associated with a concept are the defining and characteristic features of the concept.

It would be instructive, then, to formulate the Smith et al. (1974) model as a network. A network (Harary, 1969) is a digraph together with a function which assigns a value to each edge of the digraph. The digraph, in this case, would need to consist only of ordered pairs of concepts and features. For example, the ordered pair (C_i, F_j) would indicate that concept C_i had feature F_j and would be represented in the network by a directed edge from C_i to F_j . The value assigned to each edge would indicate the importance of a given feature to the definition of a concept, and, based on some threshold value, the features associated with a concept could be partitioned into defining and characteristic features. Given this definition, it is a trivial exercise to redraw, for example, Figure 1 of Smith et al. (1974, p. 216) as a network.

The difficulties that Smith et al. (1974) see confronting network models are, by Smith et al.'s own argument, "due to their failure to include structural assumptions that would make category membership a matter of degree" (p. 221). Since it is quite natural to formulate the Smith et al. (1974) model itself as a network, such criticism cannot be directed at network models in general but possibly can be leveled, as they contend, at specific network models (e.g., Quillian, 1969), although there is an argument (Collins & Loftus, Note 1) that even here the criticism is misapplied, resulting from a misinterpretation of Quillian's (1969) theory. More to the point, the argument against network models seems to be one which might better be put against digraph models in which no values are assigned to the edges. Even here, though, one could contend that it would be possible to represent defining and characteristic features within a digraph by labeling the appropriate edges as defining or characteristic. But in this case, if one is to use the terms in their mathematical sense, technically one either has two digraphs, one imposed by the defining relation and another imposed by the characteristic relation, or a network in which the function assigns nonnumerical values (for a discussion of this latter, see Harary, Norman, & Cartwright, 1965, p. 363).

A more general question is whether set-theoretic and network models are isomorphic. In the case of current psychological models, I contend the answer is affirmative. One can map any set-theoretical model onto a network model by following the procedure exemplified with the Smith et al. (1974) model. The sets associated with a given concept need only be transformed into a collection of ordered pairs in which the concept serves as the first element and the elements of the sets associated with the concept serve as second elements. One then has a relation, specified by the collection of ordered

Requests for reprints should be sent to James D. Hollan, Department of Social Sciences, Clarkson College of Technology, Potsdam, New York 13676.

pairs, and any relation can be represented by a digraph. The function associated with the digraph would assign a value or identifier to each edge to indicate which original set it was derived from. By reversing this procedure, one can map network models onto set-theoretic models. It is important to note that this argument is not meant to apply to networks as the term is used in computational linguistics. These networks are models not of human memory organization but of natural language grammars. They have evolved from finite state transition graphs in which nodes correspond to states in a finite state machine. An example of such a network is Woods' (1970) recursive transition network which is equivalent, according to Winograd (1972), to grammars expressed as programs such as PROGRAMMAR (Winograd, 1972). In these network models there are many aspects (e.g., recursion, conditional arcs, embedding) which, as far as I can see, could not be represented in a set-theoretic model.

It remains to be shown that a network formulation of the feature model can serve as an adequate structural basis for the Smith et al. (1974) two-stage feature comparison model of semantic categorization tasks and thus support the assertion that the network formulation has equal explicatory power. In the processing of a pair of text items, the underlying structural representation, whether set-theoretic or network, is accessed only at one point: In Stage 1 the features associated with the items are retrieved. It is obvious that the retrieval of the features associated with a pair of concepts in a network is no problem. As each edge directed from a concept node is traversed, the value of the edge and the feature to which it is connected is obtained.

Smith et al. (1974) make a strong case for the importance of distinguishing between defining and characteristic features of concepts and for their process model of semantic categorization. It is evident that their model can be formulated either as a set-theoretic or as a network model. There are a number of reasons, though, that one might prefer to structurally formulate their model as a network. Chief among these are the ability to then make use of the substantial literature of graph theory, an area of mathematics concerned with the abstract idea of structure (e.g., traversal algorithms such as Quillian's, 1969, breadth-first search or Bower's, 1972, depth-first search); and the ease (because of the nature

of digital computers and the availability of list-processing languages) of implementing graph-theoretical models as computer programs, a practice both common and valuable (Anderson & Bower, 1973; Hollan, 1974; Quillian, 1969; Rumelhart, Lindsay, & Norman, 1972). This practice is valuable not only because of the high degree of precision that such implementation necessitates but also because it increases the likelihood of fruitful interchange (e.g., Quillian, 1969; Winograd, 1972) between cognitive psychology and artificial intelligence, two disciplines currently very much concerned with the representation of knowledge.

REFERENCE NOTE

1. Collins, A. M., & Loftus, E. F. *A spreading activation theory of semantic processing* (Report No. 2711). Cambridge, Mass.: Bolt Beranek and Newman, 1973.

REFERENCES

- Anderson, J. R., & Bower, G. H. *Human associative memory*. Washington, D. C.: V. H. Winston, 1973.
- Bower, G. H. A selective review of organizational factors in memory. In E. Tulving & W. Donaldson (Eds.), *Organization and memory*. New York: Academic Press, 1972.
- Harary, F., Norman, R. Z., & Cartwright, D. *Structural models: An introduction to the theory of directed graphs*. New York: Wiley, 1965.
- Harary, F. *Graph theory*. Reading, Mass.: Addison-Wesley, 1969.
- Hollan, J. D. Some effects of epistemological structure on memory. *Memory and Cognition*, 1974, 2, 670-676.
- Quillian, M. R. The teachable language comprehender. *Communications of the ACM*, 1969, 12, 459-476.
- Rips, L. J., Shoben, E. J., & Smith, E. E. Semantic distance and the verification of semantic relations. *Journal of Verbal Learning and Verbal Behavior*, 1973, 12, 1-20.
- Rumelhart, D. E., Lindsay, P. H., & Norman, D. A. A process model for long-term memory. In E. Tulving & W. Donaldson (Eds.), *Organization and memory*. New York: Academic Press, 1972.
- Smith, E. E., Shoben, E. J., & Rips, L. J. Structure and process in semantic memory: A feature model for semantic decisions. *Psychological Review*, 1974, 81, 214-241.
- Winograd, T. *Understanding natural language*. New York: Academic Press, 1972.
- Woods, W. A. Transition network grammars for natural language analysis. *Communications of the ACM*, 1970, 13, 591-606.

(Received June 12, 1974)

Set-Theoretic and Network Models Reconsidered: A Comment on Hollan's "Features and Semantic Memory"

Lance J. Rips
*Department of Behavioral Sciences,
University of Chicago*

Edward E. Smith
Stanford University

Edward J. Shoben
University of Illinois

The choice of a network or set-theoretic representation correlates with two substantive differences among models of semantic memory. First, in explaining certain reaction time effects, networks models emphasize retrieval processes, while set-theoretic models often emphasize comparison processes. Second, set-theoretic models posit more semantic analysis during sentence verification than do network models. Hollan's recent demonstration of an isomorphism between set-theoretic and network representations cannot resolve these important issues.

In an earlier article (Rips, Shoben, & Smith, 1973), we characterized theories of semantic memory by means of a distinction between set-theoretic and network models. Set-theoretic models (e.g., Meyer, 1970; Schaeffer & Wallace, 1970) were those in which the meanings of words were represented by sets of semantic elements, while network models (e.g., Collins & Quillian, 1969; Rumelhart, Lindsay, & Norman, 1972) represented meanings by nodes interconnected in a network of labeled relations. Since formulating this distinction, we have witnessed two developments that appear to undermine it. We noted one of these in a recent article (Smith, Shoben, & Rips, 1974) when we mentioned the development of hybrid models which contain both network and set-theoretic assumptions (e.g., Collins & Quillian, 1972). The other challenge to our "set/net" distinction is Hollan's (1975) argument that set-theoretic representations can be translated into networks. Specifically, Hollan pointed out that a set could be equivalently represented by nodes (corresponding to the elements of that set), each connected to a common node.

Although we accept Hollan's argument, we think it essential to point out the following: The choice of one type of representation rather than another can lead to substantive processing dif-

ferences between theories, even though the representations themselves are isomorphic at some level. (See Newell & Simon, 1972, chap. 3, for a striking example of this possibility.) It seems to us that this has been the case for semantic memory models.

Thus, network and set models differ in the type of process they assume to be responsible for certain reaction time effects (Rips et al., 1973). The effects in question occur in a paradigm where one must verify statements of the form *An S is a P* (e.g., *An apple is a fruit*), and are such that verification time depends on the similarity of the subject-predicate pair. Current network models assume that these effects can be attributed primarily to retrieval of pathways between concept nodes; in contrast, set-theoretic models like those of Schaeffer and Wallace (1970) and Smith et al. (1974) suppose that these same effects are determined by comparisons of semantic elements.¹ This difference cannot be resolved by showing that set models can be translated into network representations. For suppose we adopt Hollan's (1975) proposal and allow semantic features to be connected by directed edges to the concepts they define. We now have a network

We are indebted to Herb Clark and Keith Holyoak for their thoughtful comments on the manuscript. This research was supported by U. S. Public Health Service Grant MH-19705.

Requests for reprints should be sent to Edward E. Smith, Department of Psychology, Stanford University, Stanford, California 94305.

¹ It should be noted that comparison processes are also required in network models, so that the relations between nodes in a network can be compared to the relations expressed in a sentence, while retrieval processes are also needed in set-theoretic models to access the sets of information. That is, the essential difference between the models is not which type of process is postulated, but which type of process is generally assumed to underlie the critical similarity effect.

representation, but we still have to decide on the process that underlies the verification effects of interest, either the retrieval of stored pathways between nodes or the comparison of the feature nodes of the predicate to those of the subject concept.

A related substantive issue concerns the amount of semantic analysis that one must perform to verify a statement. In network models, like that of Collins and Quillian (1969, 1972), it is sometimes possible to confirm a sentence like *An apple is a fruit* by retrieving a single link in memory. In contrast, set-theoretic models emphasize a fuller expansion of concepts in the course of sentence verification. For example, in the model of Smith et al. (1974), subject and predicate concepts are always expanded into sets of semantic features before any further processing is done. Again we have a substantive difference between existent set-theoretic and network models that is not elucidated by translating the former into the latter. For even if we do this, and construct a network representation of the Smith et al. model as Hollan does, we must still come to grips with the issue of how much semantic expansion to posit.

These two issues seem to us to be important ones in present theories of semantic memory, and it is clear that sides taken on these issues were encouraged by the choice of a network or set representation. Partly this is due to a certain type of representations being more amenable to certain types of processes. But apart from these substantive matters, there seems to be little reason to prefer networks to sets, except as a matter of theoretical convenience. While Hollan (1975) cites the literature in graph theory as a reason to prefer networks, work on set theory seems at least as adequate an underpinning for semantic memory models. Indeed, some of the best-developed models within formal semantics depend on a set-theoretic framework (e.g., Montague, 1973; Suppes, 1974). Also, implementation in terms of computer simulations can sometimes be as simple for set theories as for network theories, because of new computer languages (Raphael & Bobrow, 1974)

that use sets as well as list structures (though network structures may still be preferable when retrieval problems are an important consideration). To sum up, then, the choice between existent network and set-theoretic models remains a matter of substantive concerns.

REFERENCES

- Collins, A. M., & Quillian, M. R. Retrieval time from semantic memory. *Journal of Verbal Learning and Verbal Behavior*, 1969, 8, 240-248.
- Collins, A. M., & Quillian, M. R. How to make a language user. In E. Tulving & W. Donaldson (Eds.), *Organization of memory*. New York: Academic Press, 1972.
- Hollan, J. D. Features and semantic memory: Set-theoretic or network model? *Psychological Review*, 1975, 82, 154-155.
- Meyer, D. E. On the representation and retrieval of stored semantic information. *Cognitive psychology*, 1970, 1, 242-299.
- Montague, R. The proper treatment of quantification in ordinary English. In J. Hintikka, J. Moravcsik, & P. Suppes (Eds.), *Approaches to natural language*. Dordrecht, The Netherlands: Reidel, 1973.
- Newell, A., & Simon, H. A. *Human problem solving*. New York: Academic Press, 1972.
- Raphael, B., & Bobrow, D. G. New programming languages for artificial intelligence research. *ACM Computing Survey*, September 1974.
- Rips, L. J., Shoben, E. J., & Smith, E. E. Semantic distance and the verification of semantic relations. *Journal of Verbal Learning and Verbal Behavior*, 1973, 12, 1-20.
- Rumelhart, D. E., Lindsay, P. H., & Norman, D. A. A process model for long-term memory. In E. Tulving & W. Donaldson (Eds.), *Organization and memory*. New York: Academic Press, 1972.
- Schaeffer, B., & Wallace, R. The comparison of word meanings. *Journal of Experimental Psychology*, 1970, 86, 144-152.
- Smith, E. E., Shoben, E. J., & Rips, L. J. Structure and process in semantic memory: A featural model for semantic decisions. *Psychological Review*, 1974, 81, 214-241.
- Suppes, P. The semantics of children's language. *American Psychologist*, 1974, 29, 103-114.

(Received October 17, 1974)

Metabolic Hormones and Regulation of Feeding: A Reply to Woods, Decke, and Vasselli

Jaak Panksepp
Bowling Green State University

Woods, Decke, and Vasselli recently proposed that the insulin-growth hormone ratio, I/GH, is the signal by which body weight and feeding are regulated. Four major criticisms of the theory are discussed: (1) The I/GH ratio does not seem to have the proper characteristics to be a regulatory signal for feeding; (2) the correlation between I/GH ratios and feeding may have been prematurely elevated to causal status; (3) available evidence does not support the existence of a direct relationship between the I/GH ratio and food intake; and (4) in terms of parsimony, all predictions of the theory can be generated from insulin levels alone. It is suggested that the role of hormones in the control of feeding is more likely to be understood in terms of modifications in cellular energetics rather than in terms of any direct signaling function.

The precise manner in which mammals maintain relatively stable body energy balance has eluded definitive identification. Recently, Woods, Decke, and Vasselli (1974) proposed a possible mechanism by which body weight and food intake may be regulated. Essentially, they suggested that the regulation of body weight (or more precisely, adipose tissue) is determined by the relative secretions of insulin (I) and growth hormone (GH). When the circulating I/GH ratio increases, body weight and food intake increase; when the ratio decreases, body weight and food intake decrease.

Although there is no doubt that both insulin and growth hormone are of cardinal importance in the proper handling of body nutrient transactions, the theoretical import of the above proposal hinges largely on the manner in which changes in the I/GH ratio came to modify feeding. Without correlated changes in food intake, changes in the I/GH ratio could have only a small effect on body weight. In this reply, it will be argued that neither rational considerations nor available empirical evidence supports the idea that a change in the I/GH ratio is the antecedent condition for the occurrence of regulatory feeding behavior. Four critical questions are posed: (1) What kind of regulatory

information could the I/GH signal mediate? (2) Are systemic changes in insulin and growth hormone causes or consequences of feeding? (3) Does available experimental evidence indicate any reliable relationship between the I/GH ratio and feeding? (4) Is the proposed theory parsimonious? The position taken herein is that the role of both insulin and growth hormone in feeding and weight control is to be understood simply in terms of the metabolic changes these hormones promote. In other words, both hormones constitute part of the regulatory outflow, rather than inflow, of systems which organize the maintenance of stable body energy balance.

THE BASAL INSULIN-GROWTH HORMONE RATIO AS A REGULATORY SIGNAL

Woods et al. (1974) state that

when body weight and the insulin-growth hormone ratio are out of alignment, regulatory eating occurs. An increased ratio will cause an animal to eat much more than its usual daily ration until a new appropriate weight is attained. If the ratio is decreased, the animal will essentially stop eating until the new lower weight is attained." (p. 28)

Essentially, Woods et al. have elevated I/GH ratio to set-point status, wherefrom it dictates the level at which body fat is to be maintained via changes in regulatory feeding. But let us consider the types of information the I/GH ratio could realistically convey. A set-point, by its conceptual nature, must remain stable in the face of ongoing physiological and behavioral changes. Thus, the first problem: Secretions of both insulin and growth hormone fluctuate

The writing of this article was supported by National Science Foundation Research Grant GB-40150. I would like to thank R. Conner for a critical review of this manuscript.

Requests for reprints should be sent to Jaak Panksepp, Department of Psychology, Bowling Green State University, Bowling Green, Ohio 43403.

predictably with feeding, yielding the unlikely situation whereby a set-point fluctuates with the process it is supposed to be guiding. Insulin is secreted primarily in response to food intake (Cahill, Herrera, Morgan, Soeldner, Steinke, Levy, Reichard, & Kipnis, 1966; Steffens, 1970), and growth hormone is normally secreted several hours after meals (Hunter, Friend, & Strong, 1966) as well as during the first period of slow wave sleep each night (VanderLaan, Parker, Rossman, & VanderLaan, 1970). Possibly because of these feeling-related variations in hormone release, Woods et al. stipulate that their ratio applies to "basal (fasting)" insulin and growth hormone levels. But even granting this restriction, problems remain. Though insulin does fall to very low stable levels in starved animals and humans (Goodner & Porte, 1972), the pulsatile release of growth hormone continues (Quabbe, Schilling, & Helge, 1966). If no momentary value of growth hormone is representative of fasting, the specification of a "basal (fasting)" I/GH ratio becomes a difficult task.

But even if it is assumed that an acceptable measure of the basal I/GH ratio exists, there would be problems in specifying how the regulatory signal would operate from moment to moment. How can the basal I/GH ratio continue to act as a set-point when it is no longer around—for instance, when the animal is no longer deprived? How does the basal I/GH ratio convey the set-point message through the nonbasal I/GH jungle? After all, the highest daily I/GH ratios do not correspond to periods of hunger, and the lowest ratios do not coincide with satiety. Since insulin is secreted in response, rather than prior to food intake (Hansen & Johansen, 1970; Steffens, 1970), the lowest daily ratio corresponds to the urge to eat, and the highest ratio to the cessation of feeding. Of course, it could be argued that such paradoxes are not relevant to the theory, since the I/GH ratio only controls "regulatory eating" not "caloric eating," but that would be unsatisfactory unless it is demonstrated why existing data require two such concepts or how they are to be empirically distinguished. Such evidence was not presented.

Conceivably the I/GH ratio could provide a stable reference set-point signal if both hormones were integrated across a considerable span of time, say a day. Such integration could readily transpire for low molecular weight steroid hormones which enter cells but not for heavy protein hormones such as insulin and

growth hormone which exert their metabolic actions at the membrane level (Cuatrecasas, 1971). Prolonged membrane binding of these hormones could possibly occur somewhere in the body, but there is no evidence that such a mechanism exists. Alternatively, insulin and growth hormone may produce some intracellular "event" which is the integrated signal. But again, no such intracellular event, aside from cumulative energetic changes, is known or proposed.

INSULIN - GROWTH HORMONE RATIO AND CAUSALITY

A relatively good correlation between the average daily I/GH ratio, feeding, and body weight may well occur under certain conditions. The critical question is which is the leader and which is led. Undoubtedly, changes in circulating insulin have to precede major body weight changes, for without insulin, nutrients cannot be deposited effectively in fat depots. However, Woods et al. (1974) implicitly posit that the ratio antedates changes in feeding. In reality, the converse probably occurs—changes in feeding produce the striking changes in hormone levels.

Sims, Danforth, Horton, Bray, Glennon, & Salans (1973) described the course of experimental obesity in men who voluntarily exceeded their normal daily caloric intakes for prolonged periods. As these individuals gained weight, their insulin levels increased, growth hormone levels decreased, and their cells became resistant to the action of insulin. When they resumed eating according to desire, hormone levels returned to normal. Reductions in the I/GH ratio have also been described for obese animals (York & Bray, 1973) and humans (Salans, Knittle, & Hirsch, 1968) during food restriction. Such data, from the vantage of the proposed theory, lead to the unlikely prediction that both enforced obesity and starvation would be self-perpetuating processes. The high I/GH ratios in experimentally obese human volunteers would tend to maintain further overfeeding, and the reduced I/GH ratios of the fasting obese individuals would reinforce reductions in feeding. Since this does not happen, it is likely that the hormone changes observed in most forms of obesity are the consequences rather than the causes of hyperphagia.

EXPERIMENTAL EVIDENCE CONTRARY TO THE PROPOSED ROLE OF THE I/GH RATIO

An argument for the significance of the I/GH ratio might still be made if exogenous modifica-

tions of the ratio always produced the predicted changes in feeding behavior. Woods et al. summarize some evidence to indicate that such is the case. The most compelling support comes from the demonstration that exogenous insulin can produce a striking hyperphagia and weight gain in some species, most notably rats. It should be emphasized, however, that in other species, goats and sheep for instance, feeding is not increased by insulin (Baile & Martin, 1971; Baile & Mayer, 1968). But even in rodents, the direct role of insulin in hyperphagia would be more compelling if experimental reductions of insulin secretion led to the opposite effect. In fact, chronic insulin insufficiency (diabetes mellitus) also produces hyperphagia (Kumaresan & Turner, 1965), and this appears to reflect real energization of feeding rather than a mere compensatory response to glucose lost in the urine (Panksepp & Nance, 1972). Acute diabetes produced with mannoheptulose also does not reduce feeding reliably (Panksepp, Tonge, & Oatley, 1972). though very high doses of anti-insulin serum can inhibit feeding, probably as a consequence of debilitating ketosis (Anderson, Kilbourn, Robinson, & Wright, 1963).

It should also be noted that diabetic hyperphagia is effectively reduced by treatment with insulin (Young & Liu, 1965). Thus a low I/GH ratio (i.e., chronic diabetes) yields more feeding than a much higher I/GH ratio (i.e., diabetic on insulin therapy). Furthermore, insulin-induced hyperphagia and weight gain can be prevented or reversed simply by concurrent administration of catabolic hormones such as dexamethasone (Panksepp et al., in press). Such exceptions to the theory of Woods et al. are predictable consequences of the metabolic effects of these hormones. In the diabetic animal, exogenous insulin again allows cells to store and use energy-yielding substrates. In animals treated with dexamethasone, the potential anabolic effect of insulin is probably counteracted by cellular insulin-resistance (Åkerblom, Martin, & Gary, 1973) and accelerated fat and protein catabolism. Since insulin antagonism is also present in individuals having high endogenous growth hormone levels (Merimee & Rabin, 1973), it is not surprising that acromegaliacs have high basal insulin levels (Luft, Cerasi, & Hamberger, 1967) and that the abnormality is corrected when growth hormone levels are lowered (Karam, Grodsky, Pavlatos, & Forsham, 1965). Thus, cited evidence in support of the theory, such as growth hormone injections counteracting the effects of insulin (Zierler &

Rabinowitz, 1963), is best explained by antagonistic metabolic effects of these hormones rather than by direct changes in I/GH ratios.

Woods et al. (1974) also endeavor to document that reduction of systemic growth hormone increases feeding and that increases in growth hormone decrease feeding. Unfortunately, most of the published evidence does not agree with such a conclusion. The hypophysectomized animal is still the best available preparation for determining how reduction of growth hormone affects physiological processes. Although considerable ambiguity rises from accompanying hormone imbalances, it is clear that the most common effect of hypophysectomy on feeding is reduction of intake (Goldman, Schnatz, Bernardis, & Frohman, 1970; Smith, 1927). Furthermore, hypophysectomy can prevent the hyperphagia and obesity which is normally produced by gold-thioglucose-induced damage of the ventromedial hypothalamus (Redding, Bowers, & Schally, 1966), reduce the obesity of genetically obese (ob/ob) mice (Herbai, 1970), and reduce the food intake of weanling rats with ventromedial hypothalamic lesions (Goldman et al., 1970). All of these reductions in feeding speak against the proposed theory.

Changes in feeding after administration of exogenous growth hormone are also contrary to expectations. Feeding is increased by growth hormone injections (Andik, Sardi, & Schmidt, 1966; Goldman et al., 1970 Pfaff, 1969) and growth hormone secreting tumors (Åkerblom, Martin, Gary, & Moscarello, 1972). Of course, the concurrent weight increases can be reconciled with the proposed theory since growth hormone increases lean body mass at the expense of adipose tissue. Still, if the I/GH ratio also controls feeding, the theory should predict that all relative increases in growth hormone would tend to decrease feeding, and that clearly is not the case.

In support of their theory, Woods et al. cite the work of York and Bray (1972) where daily injections of growth hormone markedly reduced weight of hypophysectomized rats with ventromedial hypothalamic lesions. However, in that same study, growth hormone increased body weight of control animals. Since ventromedial hypothalamic lesions increase insulin secretion, growth hormone injections should have reduced the I/GH ratio least in lesioned animals, and accordingly, the controls should have lost the most weight. This did not happen. Moreover, feeding and body weight can also be increased by growth hormone in hypophysectomized

weanling rats with medial hypothalamic damage (Goldman et al., 1970).

In summary, no consistent support exists for the thesis that high I/GH ratios increase and that low ratios decrease feeding behavior of normal animals. In fact, most of the evidence summarized indicates that if a regulatory theory of feeding behavior were to be based on an I/GH ratio signal, it would probably operate in just the opposite fashion to that proposed by Woods et al. (1974). For instance, most of the reviewed data could be explained consistently if a genetically determined set-point reference signal in the hypothalamus reflected a desired I/GH level, and that the actual circulating level served as the error signal for regulatory control of feeding. If the level were higher than the prescribed set-point level (as would occur just after eating), the animal should stop eating. Conversely, were it lower (as might occur several hours after a meal), the animal would tend to eat. But until there exists direct experimental evidence for any such mechanisms, all theories of this nature will be academic.

Presently the simplest and most reasonable hypothesis concerning the normal role of insulin in feeding behavior is that it mediates satiety processes. Normally, insulin secretion occurs after feeding and allows absorbed nutrients access to intracellular metabolic machinery. Only in pathology might insulin begin to lose its satiating action by depleting blood nutrients too rapidly, thus hastening the passage of metabolic satiety, or because insulin receptors in membranes have become insensitive to the action of circulating insulin.

PARSIMONY

Finally, it seems that the proposed theory is not theoretically parsimonious. The evidence cited by Woods et al. (1974) does not require inclusion of both hormones for generating their predictions. Insulin would do alone; even growth hormone might. Since levels of the two hormones are usually reciprocally related in the body, very little additional predictive value is gained by a combination of the two. Of course reality is more complex than most theories, and the accurate depiction of the underlying physiological machinery may require equally complex explanatory constructs. But that is an empirical question, and no evidence was presented which unequivocally necessitates formation of a combinatorial insulin-growth hormone signal.

But it should be emphasized that an inclusive theory of body fat regulation (i.e., in health as well as pathology) is unlikely to be formu-

lated on insulin alone. Although it is widely recognized that basal insulin levels are directly related to the degree of obesity (Björntorp, Bergman, Varnauskas, & Lindholm, 1969; Rabinowitz, 1970), it is generally believed that insulin is, at best, permissive in the genesis of excess fat deposition. That insulin levels are not pathologically causal is indicated by experiments where insulin levels, body weight, and feeding are dissociated. Rapid weight gain can be produced with high fat diets with no increase in insulin secretion (Friedman, 1972; Lemonnier, Note 1). Adrenalectomy markedly reduces insulin levels of genetically obese mice while having little effect on their terminal level of obesity (Naeser, 1973). Obesities characterized by increased number of adipocytes, as opposed to enlargement of fat cells, are usually accompanied by normal insulin levels (Björntorp, 1972). Additional dissociations between feeding, insulin, and growth hormone levels and adiposity have been reported in weanling rats with various combinations of hypothalamic lesions and diabetes (Bernardis, Chlouverakis, Schnatz, & Frohman, 1974; Goldman et al., 1970; Goldman, Schnatz, Bernardis, & Frohman, 1972).

CONCLUSION

There is no question that body weight and feeding behavior will be modified by hormonal changes and that insulin and growth hormone will be among the main participants. Also, there is no doubt that levels of these and other hormones are modified in many forms of weight regulation disorders (see Bray and York, 1971). It is even likely that high endogenous insulin levels contribute to the weight gains and possibly to the overeating of some obesities. These facts and ideas have been around for a decade. Woods et al. (1974) have provided us with a fine summary of some of the recent work related to these issues, but until data directly implicates certain processes as being causal, there is little advantage to reifying selected correlations observed in the pathologically obese, or in assuming that such changes reflect the nature of normal regulatory mechanisms. As pointed out recently by Bray, York, and Swerloff (1973), obesity can be positively correlated with adrenal, pituitary, ovary, uterus, and liver weight—even when the obesity is prevented by underfeeding. One might reason that these changes may be even more basic to the etiology of obesity than an elevated I/GH ratio—after all these factors remain stable when a fat animal's body weight is forcibly reduced.

Correlations derived from pathological cases may obscure or caricature as well as sometimes clarify the normal functions of hormones in the body. Many hormones have biphasic effects as functions of both dosage and time. For instance, low levels of glucocorticoids can stimulate feeding (Grossie & Turner, 1965) while high levels produce profound decrements in weight and food intake (Stevenson & Franklin, 1970). Injections of insulin which eventually stimulate feeding initially suppress feeding (Lovett & Booth, 1970).

It presently remains likely that both insulin and growth hormone modulate feeding solely via their anabolic and catabolic influences on intracellular energy metabolism. The energy state of critical cells, probably in the hypothalamus, determines when feeding and subsequent spurts of hormone secretion should occur. Furthermore, if it is the energy state of certain cells that acts as the error signal, then many of the metabolic influences which act on those cells would come to bear on the strength and direction of the resultant regulatory signal. Though insulin and growth hormone would unquestionably be among the major contributors to such a control process, they would not be the only ones, and their influence would probably be understood with reference to their effects on cellular energetics rather than respective blood levels. No doubt there are many situations where hormone levels will fail to adequately reflect cellular energetics. For instance, in many forms of obesity, cells exhibit a marked resistance to the normal action of insulin (Bray & York, 1971), which may reflect decreases in the number of membrane insulin receptors (Goldfine, Kahn, Neville, Roth, Garrison, & Bates, 1973) or faulty transmission of the insulin signal to intracellular metabolism (Livingston, Cuatrecasas, & Lockwood, 1972). These changes themselves may be the consequence of pathological secretion of glucocorticoids and growth hormone (Emmer, Gorden, & Ro, 1971). In such conditions, insulin secretion reaches abnormally high levels—probably reflecting a homeostatic endeavor of the system to increase transmembrane transport of nutrients. It is likely that the increased food intake shown by organisms with such pathology is simply another manifestation of the same homeostatic endeavor. This may well lead to good correlations between insulin levels, I/GH ratios, body weight, food intake, and probably a multitude of other factors. But in the absence of any compelling evidence, the proposal of Wood et al., suggesting that such changes in

feeding are a direct consequence of one of the correlated parameters of body energetics (i.e., changes in the I/GH ratio) rather than a basic derangement of ongoing energy exchange, is gratuitous. Though pathology can often reveal the underlying principles of control systems, one must cautiously sift the causes of the pathology from the consequences.

Clearly, hormones which are important participants in regulatory processes, need not in themselves be meaningful regulatory signals. Although it is possible that some hormone(s) will eventually be found to act as unique signals to central nervous system regulatory mechanisms, until so demonstrated, it might be well to continue along traditional paths of seeking the role of hormones in feeding among the intracellular metabolic processes which are promoted or hindered.

REFERENCE NOTE

1. Lemonnier, D. *Obesity by a high-fat diet: An obesity without hyperinsulinism effect of a high-fat diet in genetically obese rats and mice.* Paper presented at the Fourth International Conference on Regulation of Food and Water Intake, Cambridge, England, August 1971.

REFERENCES

- Åkerblom, H. K., Martin, J. M., & Gary, G. L. Relative role of cortisone and growth hormone on glucose intolerance and insulin secretion in rats. *Hormones and Metabolic Research*, 1973, 5, 34-37.
- Åkerblom, H. K., Martin, J. M., Gary, G. L., & Moscarello, M. Experimental hypersomatotropism II. Metabolic effects in rats bearing the MtT-W15 tumor. *Hormones and Metabolic Research*, 1972, 4, 15-21.
- Anderson, J. W., Kilbourn, K. C., Robinson, J., & Wright, P. H. Diabetic acidosis in rats treated with anti-insulin serum. *Clinical Science*, 1963, 24, 417-430.
- Andik, I., Sardi, F., & Schmidt, P. The effect of growth hormone on feed intake and food selection. *Acta Physiologica Academiae Scientiarum Hungaricae Tomus*, 1966, 29, 177-182.
- Baile, C. A., & Martin, F. H. Hormones and amino acids as possible factors in the control of hunger and satiety in sheep. *Journal of Dairy Science*, 1971, 54, 897-905.
- Baile, C. A., & Mayer, J. Effects of insulin-induced hypoglycemia and hypoacetoemia on eating behavior in goats. *Journal of Dairy Science*, 1968, 51, 1495-1499.
- Bernardis, L. L., Chlouverakis, C., Schnatz, J. D., & Frohman, L. A. Effect of dorsomedial hypothalamic lesions before and after placement of obesity-producing ventromedial hypothalamic lesions in the weanling male rat. *Brain Research*, 1974, 69, 67-75.

- Björntorp, P. Disturbances in the regulation of food intake. *Advances in Psychosomatic Medicine*, 1972, 7, 116-147.
- Björntorp, P., Bergman, H., Varnauskas, E., & Lindholm, B. Lipid mobilization in relation to body composition in man. *Metabolism*, 1969, 18, 840-848.
- Bray, G. A., & York, D. A. Genetically transmitted obesity in rodents. *Physiological Reviews*, 1971, 51, 598-646.
- Bray, G. A., York, D. A., & Swerloff, R. S. Genetic obesity in rats. I. The effects of food restriction on body composition and hypothalamic function. *Metabolism*, 1973, 22, 435-442.
- Cahill, G. R., Jr., Herrera, M. G., Morgan, A. P., Soeldner, J. S., Steinke, J., Levy, P. L., Reichard, G. A., Jr., & Kipnis, D. M. Hormone-fuel interrelationships during fasting. *Journal of Clinical Investigations*, 1966, 45, 1751-1769.
- Cuatrecasas, P. Insulin-receptor interactions in adipose tissue cells: Direct measurement and properties. *Proceedings of the National Academy of Sciences*, 1971, 68, 1264-1268.
- Emmer, M., Gorden, P., & Ro, J. Diabetes in association with other endocrine disorders. *Medical Clinics of North America*, 1971, 55, 1057-1064.
- Friedman, M. I. Effects of alloxan diabetes on hypothalamic hyperphagia and obesity. *American Journal of Physiology*, 1972, 222, 174-178.
- Goldfine, I. E., Kahn, C. R., Neville, D. M., Jr., Roth, J., Garrison, M. M., & Bates, R. W. Decreased binding of insulin to its receptors in rats with hormone induced insulin resistance. *Biochemical and Biophysical Research Communications*, 1973, 53, 852-857.
- Goldman, J. K., Schnatz, J. D., Bernardis, L. L., & Frohman, L. A. Adipose tissue metabolism of weanling rats after destruction of ventromedial hypothalamic nuclei: Effect of hypophysectomy and growth hormone. *Metabolism*, 1970, 19, 995-1005.
- Goldman, J. K., Schnatz, J. D., Bernardis, L. L., & Frohman, L. A. Effects of ventromedial hypothalamic destruction in rats with preexisting streptozotocin-induced diabetes. *Metabolism*, 1972, 21, 132-136.
- Goodner, C. J., & Porte, D., Jr. Determinants of basal islet secretion in man. In R. O. Greep et al. (Eds.), *Handbook of physiology*, Sect. 7: *Endocrinology*, Vol. 1: *Endocrine pancreas*. Washington, D. C.: American Physiological Society, 1972.
- Grossie, J., & Turner, C. W. Effect of thyroxine, hydrocortisone and growth hormone on food intake in rats. *Proceedings of the Society for Experimental Biology and Medicine*, 1965, 118, 28-30.
- Hansen, A. P., & Johansen, K. Diurnal patterns of blood glucose, serum free fatty acids, insulin, glucagon and growth hormone in normals and juvenile diabetics. *Diabetologia*, 1970, 6, 27-33.
- Herbai, G. Weight loss in obese-hyperglycaemic and normal mice following tranauricular hypophysectomy by a modified technique. *Acta Endocrinologica*, 1970, 65, 712-722.
- Hunter, W. M., Friend, J. A. R., & Strong, J. A. The diurnal pattern of plasma growth hormone concentration in adults. *Journal of Endocrinology*, 1966, 34, 139-146.
- Karam, J. H., Grodsky, G. M., Pavlatos, F. C., & Forsham, P. A. Critical factors in excessive serum insulin response to glucose. Obesity in maturity-onset diabetes and growth hormone in acromegaly. *Lancet*, 1965, 1, 286-289.
- Kumaresan, P., & Turner, C. W. Effect of alloxan on feed consumption in rats. *Proceedings of the Society for Experimental Biology and Medicine*, 1965, 119, 400-402.
- Livingston, J. N., Cuatrecasas, P., & Lockwood, D. H. Insulin insensitivity of large fat cells. *Science*, 1972, 177, 626-628.
- Lovett, D., & Booth, D. A. Four effects of exogenous insulin on food intake. *Quarterly Journal of Experimental Psychology*, 1970, 22, 406-419.
- Luft, E. R., Cerasi, E., & Hamberger, C. A. Studies on the pathogenesis of diabetes in acromegaly. *Acta Endocrinologica*, 1967, 56, 593-607.
- Merimee, T. J., & Rabin, D. A survey of growth hormone secretion and action. *Metabolism*, 1973, 22, 1235-1250.
- Naeser, P. Effects of adrenalectomy on the obese-hyperglycemic syndrome in mice (Gene Symbol ob). *Diabetologia*, 1973, 9, 376-379.
- Panksepp, J., & Nance, D. M. Insulin, glucose and hypothalamic regulation of feeding. *Physiology and Behavior*, 1972, 9, 447-451.
- Panksepp, J., Pollack, A., Krost, K., Meeker, R., & Ritter, M. Feeding in response to protamine zinc insulin injection. *Physiology and Behavior*, 1975, in press.
- Panksepp, J., Tonge, D., & Oatley, K. Insulin and glucostatic control of feeding. *Journal of Comparative and Physiological Psychology*, 1972, 78, 226-232.
- Pfaff, D. Sex differences in food intake changes following pituitary growth hormone or prolactin injections. *Proceedings of the 77th Annual Convention of the American Psychological Association*, 1969, 4, 211-212.
- Quabbe, H. J., Schilling, E., & Helge, H. Pattern of growth hormone secretion during a 24-hour fast in normal adults. *Journal of Clinical Endocrinology and Metabolism*, 1966, 26, 1173-1177.
- Rabinowitz, D. Some endocrine and metabolic aspects of obesity. *Annual Review of Medicine*, 1970, 21, 241-258.
- Redding, T. W., Bowers, C. Y., & Schally, A. V. Effects of hypophysectomy on hypothalamic obesity in CBA mice. *Proceedings of the Society for Experimental Biology and Medicine*, 1966, 121, 727-729.
- Salans, L. B., Knittle, J. L., & Hirsch, J. The role of adipose cell size and adipose tissue insulin sensitivity in the carbohydrate intolerance of human obesity. *Journal of Clinical Investigations*, 1968, 47, 153-165.

- Sims, E. A. H., Danforth, E., Jr., Horton, E. S., Bray, G. A., Glennon, J. A., & Salans, L. B. Endocrine and metabolic effects of experimental obesity in man. In R. D. Greep (Ed.), *Recent Progress in Hormone Research* (Vol. 29). New York: Academic Press, 1973.
- Smith, P. E. The disabilities caused by hypophysectomy and their repair: The tuberal (hypothalamic) syndrome in the rat. *Journal of the American Medical Association*, 1927, 88, 158-161.
- Steffens, A. B. Plasma insulin content in relation to blood glucose level and mean pattern in the normal and hypothalamic hyperphagic rat. *Physiology and Behavior*, 1970, 5, 147-151.
- Stevenson, J. A. F., & Franklin, C. Effects of ACTH and corticosteroids in the regulation of food and water intake. In D. de Wied & J. A. W. M. Weinen (Eds.), *Progress in Brain Research* (Vol. 32). Amsterdam: Elsevier, 1970.
- VanderLaan, W. P., Parker, D. C., Rossman, L. G., & VanderLaan, E. F. Implications of growth hormone release in sleep. *Metabolism*, 1970, 19, 891-897.
- Woods, S. C., Decke, E., & Vasselli, J. R. Metabolic hormones and regulation of body weight. *Psychological Review*, 1974, 81, 26-43.
- York, D. A., & Bray, G. A. Dependence of hypothalamic obesity on the pituitary and the adrenal gland. *Endocrinology*, 1972, 90, 885-894.
- York, D. A., & Bray, G. A. Genetic obesity in rats. II. The effect of food restriction on the metabolism of adipose tissue. *Metabolism*, 1973, 22, 443-454.
- Young, T. K., & Liu, A. C. Hyperphagia, insulin and obesity. *Chinese Journal of Physiology*, 1965, 19, 247-253.
- Zierler, K. L., & Rabinowitz, D. Roles of insulin and growth hormone, based on studies of forearm metabolism in man. *Medicine*, 1963, 42, 385-402.

(Received February 28, 1974; revision
received June 24, 1974)

Insulin, Growth Hormone, Body Weight, and Feeding: A Reply to Panksepp

Stephen C. Woods
University of Washington

Elisabeth Kaestner and Joseph R. Vasselli
Columbia University

Panksepp has criticized our recent theory on the regulation of body weight by mammals. This note constitutes a rebuttal. It is shown that the papers cited by Panksepp actually agree with the theory and that his criticisms arose out of misinterpretation rather than faulty logic on our part.

In a reply to our article on "Metabolic Hormones and the Regulation of Body Weight" (Woods, Decke, & Vasselli, 1974), Panksepp (1975) challenged our theory of body weight regulation on a number of grounds. We welcome the criticisms for they give us the opportunity to clear up some potential areas of general misinterpretation.

In brief, our theory is based upon the high correlation which exists between body weight¹ and the insulin-growth hormone ratio (I/GH ratio) of mammals (for original citations, see Woods et al., 1974, or Woods and Porte, 1974). A high I/GH ratio corresponds to a relative excess of body fat and vice versa. The normal or "ideal" ratio is assumed to be genetically determined and to prevail under lack of external challenges to the system. Presumably, some area of the central nervous system, probably in the hypothalamus, monitors the I/GH ratio and can induce variations in activity to pathways influencing the rates of secretion of both insulin and growth hormone (and thereby can control body weight). It is also assumed that a negative feedback system exists in which externally induced deviations of the I/GH ratio

from its ideal value will cause a compensatory shift in the centrally controlled release of insulin and growth hormone.

Deviations of the I/GH ratio from its ideal value can be induced exogenously in three ways: (a) by direct manipulation of the levels of the hormones; (b) by manipulations affecting the central regulation of hormone secretion (i.e., alteration of central nervous system activity); and (c) by manipulations affecting the peripheral stimulation of hormone secretion (e.g., by variation in the amount of food consumed). If such changes of the value of the ratio cannot be compensated by the negative feedback system, the new prevailing I/GH ratio will dictate a new body weight. Thus, when body weight and the ratio get out of alignment, "regulatory eating" will occur, causing body weight to adjust to the prevailing ratio. When the exogenous influences on the ratio are removed, body weight will revert to normal.

We will now turn to the four specific questions raised by Panksepp. (1) "What kind of regulatory information could the I/GH signal mediate?" (Panksepp, 1975, p. 158). Panksepp correctly points out that it is difficult to reconcile the fact that only the basal (fasting) I/GH ratio corresponds to body weight and that basal conditions may never normally occur in mammals; that is, plasma levels of both insulin and growth hormone covary with a number of behavioral events including meals, exercise, stress, and sleep. On the other hand, Panksepp (1975) does concede that "the I/GH ratio could provide a stable reference set-point signal if both hormones were integrated across a considerable span of time, say a day" (p. 159). This apparent problem would therefore be solved if the I/GH ratio could be dissociated from ongoing behavioral events, at least as it affects the central nervous system. The fact that such a system has not been identified does not invalidate the theory, particularly since

The writing of this paper was supported by Research Grant AM-17112 from the National Institute of Health. The authors would like to thank P. J.ulkosky, R. H. Lovely, and D. Porte, Jr., for their helpful comments.

Requests for reprints should be sent to Stephen C. Woods, Department of Psychology NI-25, University of Washington, Seattle, Washington 98195.

Elisabeth Kaestner (formerly Decke) is now at the New York State Drug Abuse Control Commission, No. 2, World Trade Center, New York 10047.

Joseph R. Vasselli is now at the Department of Psychology, Brooklyn College, Brooklyn, New York.

The term "body weight" refers to body fat rather than to lean body mass throughout this

there is no contrary data. In fact, evidence for insulin-sensitive sites in the central nervous system is becoming impressive (see Woods and Porte, 1974, for a review). (2) "Are systemic changes in insulin and growth hormone causes or consequences of feeding?" (Panksepp, 1975, p. 158). There is no doubt that food intake per se can change the I/GH ratio. In fact, we explicitly stated that "changes of the ratio can be induced by . . . forced variation in the amount of food consumed" (Woods et al., 1974, p. 27). It should be emphasized that our theory of body weight regulation predicts only those episodes of eating which reflect discrepancies between body weight and the I/GH ratio (i.e., regulatory eating). Since most episodes of food intake, particularly in normal animals (i.e., "caloric eating"), are *not* triggered by such discrepancies—but rather are external to the proposed system—they were not strongly considered. Panksepp repeatedly failed to make this distinction. It is well known that a variety of conditions, including environmental temperature, exercise, palatability of diet, and social conditions of eating, to name a few, can influence food intake and, depending upon their intensity and duration, can present challenges to the ideal (or prevailing) I/GH ratio. Any induced changes of the ratio, however, will be maintained only as long as the external conditions which alter feeding prevail. The fact that experimentally overfed obese subjects (e.g., Sims, Horton, & Salans, 1971) easily revert to their normal body weight (as well as the fact that naturally obese individuals easily revert to their high body weight) once allowed to free feed is not contradictory to our theory as Panksepp claimed. Our theory predicts that excessive food intake leads to an increase of the I/GH ratio and causes body weight (fat) to increase. This process, however, is not self-perpetuating as Panksepp claimed since the ratio will be increased only as long as the excessive food intake is maintained. Once this external influence on insulin and growth hormone secretion is removed, the I/GH ratio declines, regulatory eating (fasting, in this instance) occurs, and the weight and the ratio are returned to the central nervous system-controlled levels. However, it seems possible that more chronic changes of feeding might eventually lead to functional adaptations of the secretory glands and consequently to more permanent deviations of the prevailing I/GH ratio from its normal value. (3) "Does available experimental evidence indicate any reliable relationship between the I/GH ratio and feeding "

(Panksepp, 1975, p. 158). The cited examples of chronic insulin insufficiency accompanied by hyperphagia are not valid as counterevidence to our theory since they are instances of caloric eating (to comply with energy requirements) rather than of regulatory eating (to bring body weight into alignment with the I/GH ratio). In addition, Panksepp failed to mention that when insulin was lowered in all of the studies he cited, body weight fell in spite of the acute hyperphagia, exactly as our theory predicts. Further, in the studies Panksepp cited in which growth hormone was administered and in which fat stores were analyzed, the fat stores actually decreased although lean body mass may have increased resulting in a slight increase in overall body weight.

A second area of contention concerns the effects of hypophysectomy upon body weight. As we pointed out (Woods et al., 1974, p. 33), removal of the pituitary as a means of reducing growth hormone is a crude measure at best because many other hormones secreted by the gland are simultaneously eliminated. The findings on animals with ventromedial hypothalamic (VMH) lesions in addition to hypophysectomy, which Panksepp cited, therefore make a weak counterargument to our theory, especially since most of the data on such rats supports our theory. The VMH lesions decrease the growth rate, especially in young rats (Bernardis & Skelton, 1966). It is therefore not surprising that VMH-lesioned hypophysectomized rats lost weight when injected with growth hormone and that hypophysectomized controls without VMH lesions gained weight (York & Bray, 1972). Most likely, the controls gained lean body mass such that the proportion of body fat decreased, as our theory predicts. Cox, Kakolewski, and Valenstein (1968) showed that VMH-lesioned hypophysectomized rats gained more weight per gram of food intake than nonhypophysectomized but VMH-lesioned controls. There is also evidence that hypophysectomized animals have a greater proportion of body fat than controls (Lee & Ayres, 1936) and that hypophysectomized rats injected with growth hormone have a smaller proportion of body fat than controls (Han, 1967).

In his article, Panksepp also stated that chronic injections of insulin do not always cause an elevation of body weight, citing two experiments on ruminants. Independent of the use of negative results as supportive evidence, it may be that ruminants, with their unique digestive systems, simply require more insulin. Finally, Panksepp himself has reached the

same conclusion that we have regarding insulin and feeding, stating that "The VMH appears to modulate feeding in response to changing levels of systemic insulin" (Panksepp & Nance, 1972, p. 451). (4) "Is the proposed theory parsimonious?" (Panksepp, 1975, p. 158). Panksepp suggested that insulin alone (without growth hormone) provides as much information as the ratio and that the observed effects of growth hormone on body weight could be simply explained by insulin antagonism of growth hormone (p. 158). We basically agree with this statement. However, we prefer the concept of an I/GH ratio because it is more easily operationalized than "functional" insulin.

In his review, Panksepp also claims that there are many instances of dissociation between the ratio and body weight, citing seven references. Of the seven, five are instances in which *eating* and insulin, *not weight* and insulin, occur in what Panksepp considers a strange relationship. It is because of situations like these that we purposely distinguished between caloric eating and regulatory eating and focused upon body weight regulation. Of the remaining two citations, one is unpublished and cannot be evaluated. The final reference (Björntorp, 1972), according to Panksepp, states that certain obese states are accompanied by normal insulin levels. A careful reading of Björntorp, however, reveals that he reached the opposite conclusion, the one predicted by our theory (Björntorp, 1972, pp. 124 & 134).²

Finally, let us return to Panksepp's (1975) two major criticisms; that is, that the theory "hinges largely on the manner in which changes of the I/GH ratio come to modify feeding" and that "neither rational considerations nor available empirical evidence supports the idea that a change in the I/GH ratio is the antecedent condition for the occurrence of regulatory feeding behavior" (p. 158). We discussed caloric eating above and it should be clear that caloric eating is not a necessary condition for changes of body weight.

Many experiments show that the I/GH ratio can alter body fat in the absence of changes of food intake (e.g., Bernardis & Frohman, 1970; Frohman & Bernardis, 1968; Han, 1967; Hustvedt & Lovo, 1972). Additionally, some more

recent research supports our conception of the control of regulatory eating. Although the evidence that a lesion of the VMH causes an increase of insulin that precedes hyperphagia and obesity was scant when the theory was first published, it is now well established. When the VMH is functionally lesioned in awake rats by use of the local application of a reversible anesthetic, insulin values are significantly elevated within 3 min in the absence of food intake, and they return to normal when the anesthetic wears off (Kulkosky, Porte, & Woods, in press). Further, two other recent studies have reported that the increase of insulin occurs when hyperphagia is prevented following VMH lesions to rats (Martin, Konijnendijk, & Bouman, 1974; Tannenbaum, Paxinos, & Bindra, 1974). Finally, Hustvedt and Lovo (1972) showed that the hyperinsulinemia that occurred in rats with VMH lesions but restricted food intake could be used to predict their subsequent gain of weight given hyperphagia. Taken together, these studies strongly imply that there is an increase of insulin following a VMH lesion and that the magnitude of the increase determines the subsequent gain of weight and duration of hyperphagia, exactly as the theory predicts. Finally, Powley and Opsahl (1974), in an elegant recent experiment, showed that the obesity following a VMH lesion is eliminated following section of the vagus nerves, the nerves which carry impulses from the central nervous system to the pancreas and influence insulin secretion (see Woods and Porte, 1974, for a review).

In conclusion, although we, as well as Panksepp (1975), feel that something as simple as the I/GH ratio cannot account for all forms of body weight regulation, it does so better than other theories. Further, the fact that the large number of citations used by Panksepp to criticize the theory actually are compatible with it suggests to us that we may be on the right track.

REFERENCES

- Bernardis, L. L., & Frohman, L. A. Effect of lesion size in the ventromedial hypothalamus on growth hormone and insulin levels in weanling rats. *Neuroendocrinology*, 1970, 6, 319-328.
- Bernardis, L. L., & Skelton, F. R. Growth and obesity following ventromedial hypothalamic lesions placed in female rats at four different ages. *Neuroendocrinology*, 1966, 1, 265-275.
- Björntorp, P. Disturbances in the regulation of food intake. *Advances in Psychosomatic Medicine*, 1972, 7, 116-147.

²This point is somewhat controversial since Björntorp (1972) hypothesizes that insulin and body weight might occasionally be out of alignment, citing unpublished data as support (e.g., pp. 123 and 135). All of the data presented by Björntorp, however, agree with our theory.

- Cox, V. C., Kakolewski, J. W., & Valenstein, E. S. Effects of ventromedial hypothalamic damage in hypophysectomized rats. *Journal of Comparative and Physiological Psychology*, 1968, 65, 145-148.
- Frohman, L. A., & Bernardis, L. L. Growth hormone and insulin levels in weanling rats with hypothalamic lesions. *Endocrinology*, 1968, 82, 1125-1132.
- Han, P. W. Hypothalamic obesity in rats without hyperphagia. *Transactions of the New York Academy of Sciences*, 1967, 30, 229-243.
- Hustvedt, B. E., & Lovo, A. Correlation between hyperinsulinemia and hyperphagia in rats with ventromedial lesions. *Acta Physiologica Scandinavica*, 1972, 84, 29-33.
- Kulkosky, P. J., Porte, D., Jr., & Woods, S. C. Intrathecal pentobarbital elevates plasma insulin of rats. *Experientia*, in press.
- Lee, M., & Ayres, G. B. The composition of weight loss and the nitrogen partition of tissues in rats after hypophysectomy. *Endocrinology*, 1936, 20, 489-495.
- Martin, J. M., Konijnendijk, W., & Bouman, P. R. Insulin and growth hormone secretion in rats with ventromedial hypothalamic lesions maintained on restricted food intake. *Diabetes*, 1974, 23, 203-208.
- Panksepp, J. Metabolic hormones and regulation of feeding: A reply to Woods, Decke, and Vasselli. *Psychological Review*, 1975, 82, 158-164.
- Panksepp, J., & Nance, D. M. Insulin, glucose and hypothalamic regulation of feeding. *Physiology and Behavior*, 1972, 9, 447-451.
- Powley, T. L., & Opsahl, C. A. Ventromedial hypothalamic obesity abolished by subdiaphragmatic vagotomy. *American Journal of Physiology*, 1974, 226, 25-33.
- Sims, E. A. H., Horton, E. S., & Salans, L. B. Inducible metabolic abnormalities during development of obesity. *Annual Review of Medicine*, 1971, 22, 235-250.
- Tannenbaum, G. A., Paxinos, G., & Bindra, D. Metabolic and endocrine aspects of the ventromedial hypothalamic syndrome in the rat. *Journal of Comparative and Physiological Psychology*, 1974, 86, 404-413.
- Woods, S. C., Decke, E., & Vasselli, J. R. Metabolic hormones and regulation of body weight. *Psychological Review*, 1974, 81, 26-43.
- Woods, S. C., & Porte, D., Jr. Neural control of the endocrine pancreas. *Physiological Reviews*, 1974, 54, 596-619.
- York, D. A., & Bray, G. A. Dependence of hypothalamic obesity on insulin, the pituitary and the adrenal gland. *Endocrinology*, 1972, 90, 885-894.

(Received August 22, 1974)

Thesaurus of Psychological Index Terms

Contains over 4,000 terms representing an authoritative, unambiguous, and hierarchically structured vocabulary used by psychology and related disciplines. This vocabulary will facilitate precise content representation of concepts described in psychological literature and allow for more flexibility in automated search and retrieval operations. These two factors promote more effective communication of psychological concepts in original works and research, ultimately aiding the content of publications to be focused in accordance with the interest patterns of psychologists.

This **Thesaurus**, developed by the Psychological Abstracts Information Services, is divided into three sections:

The **relationship section** shows the relationships among concepts that are associated with each other. Term relationships noted are: synonymous, broader, narrower, or related.

The **alphabetical section** presents each preferred term in alphabetical order, enabling rapid identification of terms used for indexing.

The **hierarchical section** contains 17 broad content areas in which terms are presented in descending order according to breadth of concept represented. The 17 categories correspond to major classifications used in *Psychological Abstracts*.

The relationship, alphabetical, and hierarchical sections will assist the user in selecting appropriate terms for different uses.

Administrators

- determine areas of research for an organization
- assure representative expertise in personnel selection

Researchers and Authors

- find synonyms for a term representing a concept
- select terms that represent a concept
- define subject matter topics for a paper or book

Educators

- structure a curriculum
- use as a tutorial aid

Consultants

- search the literature
- aid in retrieval of information

Librarians

- provide access to psychological literature
- aid in cataloging psychology material
- aid in ordering new psychological literature

Students

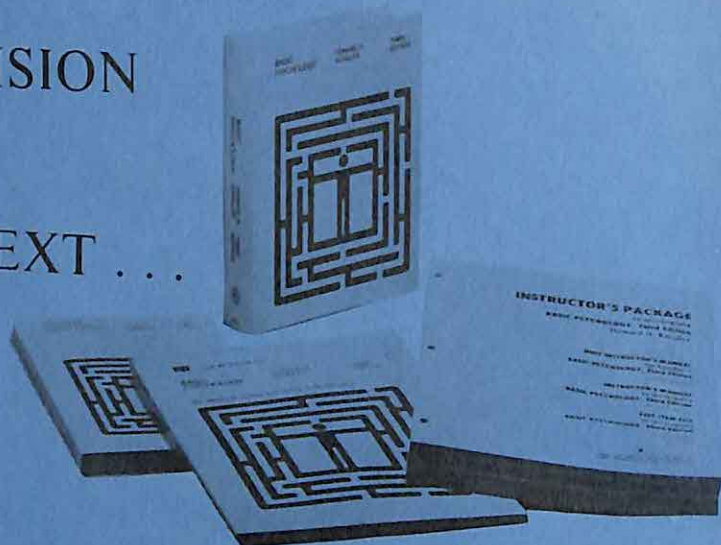
- locate concepts relevant to psychology
- focus areas of study

The **Thesaurus** contains 350 pages with a hard-cover and costs \$12. Please note that all orders amounting to \$15 or less must be accompanied by full payment except for those orders submitted on institutional purchase order forms.

American Psychological Association
Publication Sales Department
1200 17th Street, N.W.
Washington, D.C. 20036



A NEW REVISION
OF A
CLASSIC TEXT . . .



BASIC PSYCHOLOGY

THIRD EDITION

By

HOWARD H. KENDLER

University of California, Santa Barbara

As with previous editions of *Basic Psychology*, this edition carefully preserves the view that psychology is a science, data bound and empirically based. Yet, exciting developments in such areas as humanistic psychology, social learning theory, ESP, contagious violence, attribution theory, cognition, and redundancy in perception are shown to be analyzable from the scientific perspective.

A clear, precise presentation gives each student a grounded understanding of psychology as a scientific endeavor as well as an appreciation for the intriguing complexities of human behavior.

In discussing social issues, Kendler is not glamorous, just reliably sound.

A complete learning and teaching aid package accompanies the text and includes:

Student Study Guide • PSI Unit Workbook • Instructor's Package



W. A. BENJAMIN, INC.
2727 Sand Hill Road, Menlo Park, CA 94025

PSYCHOLOGICAL REVIEW

GEORGE MANDLER, EDITOR

JEAN M. MANDLER, ASSOCIATE EDITOR

UNIVERSITY OF CALIFORNIA, SAN DIEGO

AKC
Docket 2
Issue 15
Librarian
R
Dad
28/8/75

Vol. 82, No. 3

May 1975

Articles

A Balance Theory Interpretation of Dissonance

Chester A. Insko, Stephen Worchel, Robert Folger, and
Arūnas Kutkus 169

Perturbation Model for Letter Identification

George Welford 184

Role of the Hypothalamus in the Regulation of Food and
Water Intake Sebastian P. Grossman 200

Editorial Consultants

The editor gratefully acknowledges the efforts of the following consultants who helped review manuscripts for this issue. They are part of a large group of psychologists who have kindly agreed to assist in the editing of manuscripts for *Psychological Review*.

JOHN R. ANDERSON	COURTNEY CAZDEN	SARAH LICHTENSTEIN	DAVID E. RUMELHART
NORMAN H. ANDERSON	BARRY E. COLLINS	JAMES MCGAUGH	MARILYN SHATZ
JACOB BECK	JAMES H. DAVIS	D. E. MEYER	RICHARD M. SHIFFRIN
GORDON H. BOWER	ROBYN DAWES	JEFFREY MILLER	SAUL STERNBERG
ROBERT M. BOYNTON	ELLIOTT ROGERS	BENNETT MURDOCK, JR.	PHILIP E. TEITELBAUM
JOHN L. BROWN	PAUL E. GOLD	ALLEN NEWELL	JAMES T. TOWNSEND
M. PHILIP BRYDEN	DOUGLAS HINTZMAN	ELISSA NEWPORT	ELLIOTT S. VALENSTEIN
ALEXANDER M.	JULIAN HOCHBERG	KENT L. NORMAN	STUART VALINS
BUCHWALD	CHESTER A. INSKO	PAUL A. OBRIST	GERALD WESTHEIMER
ROSALIND D.	MARCEL KINSBOURNE	ELEANOR ROSCH	WAYNE A. WICKELGREN
CARTWRIGHT	DAVID H. KRANTZ	ERNST ROTHKOPF	THOMAS D. WICKENS

The *Psychological Review* publishes articles that make theoretical contributions to any area of scientific psychology. Preference is given to papers that advance theory rather than review it, and to statements that are specifically theoretical rather than programmatic. Surveys of the literature, problems of method and design, and papers that primarily report empirical findings are ordinarily not appropriate. Letters not exceeding 250 words on topics of theoretical interest should be submitted in duplicate.

Preparation of Manuscripts: Instructions for the preparation of manuscripts appear in the *Publication Manual of the American Psychological Association* (2nd ed.). Submit three copies; one copy must be the original typed copy. All parts of the manuscript must be double spaced. In order to facilitate blind reviewing, the author's name should not appear on the manuscript or abstract, but each copy of the manuscript should include a separate title page showing the title and the author's name and affiliation. Footnotes that identify the author should be typed on separate pages. Figures should be submitted as glossy prints ready for reproduction; duplicate figures may be photocopies. Metric equivalents must be given for all nonmetric measurements. Authors should keep a copy of the manuscript to guard against loss.

Abstracts: Manuscripts must be accompanied by an abstract of 100 to 175 words typed on a separate sheet of paper. Detailed instructions for preparation of abstracts appear in the *Publication Manual* or may be obtained from the Editor.

Manuscripts: Submit manuscripts in triplicate to the Editor, GEORGE MANDLER, Department of Psychology, University of California, San Diego, La Jolla, California 92037 according to instructions provided above.

Copyright and Permissions: All rights reserved. Written permission must be obtained from the American Psychological Association for copying or reprinting text of more than 500 words, tables, or figures. Permissions are normally granted contingent upon like permission of the author, inclusion of the APA copyright notice on the first page of reproduced material, and payment of a fee of \$10 per page of APA material. Permission fees are waived for authors who wish to reproduce their own articles. Address requests for permission to the Permissions Office, American Psychological Association, 1200 Seventeenth Street, N.W., Washington, D. C. 20036.

Subscriptions: Subscriptions are available on a calendar year basis only (January through December). Nonmember rates for 1975: \$15 domestic, \$16 foreign, \$3 single issue. APA member rate: \$6. Write to Subscription Section, APA.

Reprints: Authors may order reprints of their articles from the printer when they receive galley proofs.

Back Issues and Back Volumes: For information regarding back issues or back volumes write to Order Department, APA, Box 1017, Greenwich, Connecticut 06830; University Microfilms, Ann Arbor, Michigan 48106; or Princeton Microfilms, Princeton, New Jersey 08540.

Change of Address: Send change of address notice and a recent mailing label to the attention of the Subscription Section, APA, 30 days prior to the actual change of address. Undelivered copies resulting from address changes will not be replaced; second-class mail is not forwarded unless subscribers notify the local post office in writing that they will guarantee second-class forwarding postage.

Published bimonthly (beginning in January) in one volume per year, six issues, by the American Psychological Association, Inc., 1200 Seventeenth Street, N.W., Washington, D. C. 20036. Printed in the U.S.A. Second-class postage paid at Washington, D. C., and additional mailing offices.

Journal Staff

ANITA DEVIVO
Executive Editor

MARY W. GRIMBALL
Supervising Technical Editor

ROBERT J. HAYWARD
Advertising Manager

ANNE REDMAN
Subscription Manager

KENDALL C. PATTERSON
Technical Editor

PSYCHOLOGICAL REVIEW

Copyright © 1975 by the American Psychological Association, Inc.

VOL. 82, No. 3

MAY 1975

A Balance Theory Interpretation of Dissonance

Chester A. Insko, Stephen Worchel, Robert Folger, and Arūnas Kutkus
University of North Carolina

An argument was developed that the crucial cognitive band in the Festinger and Carlsmith paradigm situation is as follows: The self (+) is responsible for (?) arguing that the task is enjoyable (-). Assuming that the self is positively evaluated and that the counterattitudinal advocacy is negative, the remaining issue becomes whether or not the subject will accept responsibility. Following an analysis of the meaning or meanings of the term *responsibility*, it was concluded that responsibility will most likely be accepted when there is both high choice and low reward. In this case, however, the above cognitive band becomes unbalanced. Such imbalance is then reduced by a change in the evaluation of the counterattitudinal act. Although such a change could be achieved by the subject deciding that it is not too bad to mislead someone, the change is more plausibly achieved by the subject deciding that "I did not mislead the other subject, since I really do believe that the task is enjoyable." The above analysis also leads to an explanation of Calder, Ross, and Insko's reinforcement effect, as well as to a prediction that a manipulation of investment of time—as opposed to reward—along with choice and consequences would produce a triple interaction. This prediction was confirmed. Finally, a balance explanation of other forced-compliance situations was briefly sketched, along with an analysis of free-choice situations.

In 1959, Festinger and Carlsmith introduced what has come to be the best known paradigm for the study of dissonance. After subjects spent an initial period working on a boring task, they were offered either a small or a large reward for telling a waiting "subject" that the task was interesting and enjoyable. After the subjects confronted and convinced the waiting subject (actually a stooge), they reported that the task was more enjoyable in the small- or low-reward condition than in the large- or high-reward condition. Thus Festinger and Carlsmith obtained results consistent with one of dissonance theory's nonobvious predictions, namely, the less the reward used to induce counterattitudinal advocacy the greater the attitude change in the direction of advocacy.

The Festinger and Carlsmith study stimulated a large number of subsequent investi-

gations. Some of the studies used the same dull-task procedure and some did not. In all of the studies, however, different levels of reward were associated with some form of counterattitudinal advocacy. The results of these studies are seemingly inconsistent. Some studies (e.g., Cohen, 1962) found the same inverse or negative relationship between reward and attitude change found by Festinger and Carlsmith. Such an inverse relation has come to be known as a dissonance effect. Other studies (e.g., Elms & Janis, 1965; Rosenberg, 1965), however, found a direct or positive relation between reward and attitude change, which has come to be known as a reinforcement effect. And still other studies (e.g., Melson, Calder, & Insko, 1969) found no relation between reward and attitude change. This large literature has been recently and exhaustively reviewed by Collins and Hoyt (1972).

TABLE 1

MEAN ENJOYABLENESS IN HIGH-CONSEQUENCES-
ACTORS CONDITION OF THE CALDER, ROSS,
AND INSKO STUDY^a

Reward	High choice	Low choice
$\frac{1}{2}$ hour	21.33	9.60
2 hours	14.80	16.20

Note. Control actors = 13.20. All ratings on a 36-point scale, with higher numbers indicating more enjoyment.

^a Adapted from Calder, Ross, and Insko (1973, p. 91).

Calder, Ross, and Insko (1973) attempted to determine the interactive factors accounting for the apparent inconsistency of results. We will describe this study in some detail, since it provides an illustrative example for our theoretical position. Calder et al. manipulated four factors in the context of the Festinger and Carlsmith (1959) paradigm. These factors were as follows: reward ($\frac{1}{2}$ hour credit or 2 hours credit), choice (high or low), consequences (high or low), and actor-observer. The latter two factors require explanation. In the low-consequences condition the stooge was unconvinced by the subject's counterattitudinal advocacy. In the high consequences condition the stooge was convinced and also stated that since the experiment was interesting and enjoyable he would stay and be a subject, rather than leave and study for a quiz as he had originally intended. The actor-observer factor was manipulated by pairing every actor-subject with an observer-subject. The observer-subject observed the interaction between the actor-subject and the waiting subject through a one-way mirror.

An analysis of variance of the judged enjoyableness of the task revealed a triple interaction: Reward $\times$ Choice $\times$ Consequences. This triple interaction resulted from a dissonance effect when there were both high consequences and high choice, and a reinforcement effect when there were high consequences and low choice. The double interaction between reward and choice within high consequences was significant. The double interaction between reward and choice within low consequences was slightly reversed in form and was non-significant. The means for the actor-subjects within the high-consequences con-

dition are presented in Table 1. These means are based on ratings of the task on a 36-point scale, in which high numbers indicate the most enjoyableness. The means for the observer-subjects revealed a similar double-interaction pattern within high consequences. Control subjects who simply worked on the dull task and then rated its enjoyableness had a mean rating of 13.20. This control mean differed significantly *only* from the mean in the high-dissonance cell, which was 21.33. Although the variables were operationalized somewhat differently, Collins and Hoyt (1972) obtained a similar triple interaction, in a study which further found that the greater amount of attitude change in the high-dissonance cell was still present on a 2-week delayed posttest.

The central thesis of the present article is that balance theory (Heider, 1946, 1958) or affective-cognitive consistency theory (Rosenberg, 1956, 1965; Rosenberg & Abelson, 1960) provides a framework that can be used to account for the above dissonance results and indeed all dissonance results. It is true, of course, that Heider, Rosenberg, and Rosenberg and Abelson did not use their balance or consistency theories in this fashion. Rather, they presented their theories as general perspectives on the resolution of imbalance in cognitive bands (i.e., two cognitively linked attitude objects). The important problem of specifying the relevant cognitive bands for particular situations is largely untreated. It is our contention that the important cognitive band for the Festinger and Carlsmith (1959) paradigm is one that involves the self: The self (+) is responsible for (?) advocating that the dull task is enjoyable (?). At this point the only sign that has been supplied is the positive one for the self. Heider, of course, explicitly assumes that pLp or that the subject holds himself in high regard. Although the assumption may not be universally true, common sense indicates that it is certainly true for the majority of people. An emphasis on the importance of the self and the self's maintenance of high self-esteem as an explanation of dissonance has been made by Aronson (1968), Nel,

Helmreich, and Aronson (1969), and Brown (1965).

What about the sign of the second object of affective significance in the cognitive band, advocating that the dull task is enjoyable? Counterattitudinal advocacy need not always be unpleasant or negative. There may indeed be instances, in a debate for example, in which counterattitudinal advocacy is fun. We believe, however, that in the Festinger and Carlsmith (1959) paradigm, counterattitudinal advocacy is more likely to be negative. After all, in this situation the subject does attempt to mislead someone, that is, to lie. The degree of negativeness, however, is probably not too great in a low-consequences condition, in which the subject was not successful in misleading the waiting subject. In a high-consequences condition, however, the degree of negativeness is obviously more extreme. This is particularly true in the Calder et al. (1973) study, in which the subject not only misled the waiting subject but also prevented him from studying for an upcoming quiz. Within a high-consequences condition, then, we will assume that the counterattitudinal advocacy is negative. The cognitive band thus changes: the self (+) is responsible for (?) advocating that the dull task is enjoyable (-). An emphasis on the importance of high or negative consequences in the production of dissonance has previously been made by Brown (1965), Aronson (1968), Nel et al. (1969), Collins (1968), and Cooper and Worchel (1970).

The final part of the cognitive band is the cognitive linkage, "is responsible for." In Heider's language this is a unit relation, in Rosenberg's language an instrumental cognition. The first emphasis on the importance of responsibility in the production of dissonance of which we are aware is contained in an unpublished paper by Cooper (Note 1). Cooper's paper is worth considering in some detail because its contents are not generally known and because the paper discusses what we consider a crucial issue. Cooper's central argument is that "responsibility is a necessary condition for the arousal of dissonance" (p. 64). The main conceptual problem that Cooper

confronts involves the relationship between responsibility and choice. Cooper argues that generally "choice is relevant to responsibility when the act of choice serves to confront an individual with the responsibility of his own behavior" (p. 31). Nonetheless, Cooper points out that it is possible to have "responsibility without choice" (p. 45) and "choice without responsibility" (p. 51). Responsibility without choice is illustrated by the individual with high self-esteem who is aware that he possesses a negative characteristic, and also by some behavior which an individual recognizes as being a product of the unconscious. Choice without responsibility occurs when an individual makes a choice without foreseeing the consequences. Cooper discusses Brehm and Cohen's (1962) well-known example of the farmer who decides to plant tobacco rather than corn and then sometime later learns that tobacco causes a deadly disease. According to Brehm and Cohen the farmer should experience dissonance, since choice and volition were present. Cooper, on the other hand, argues that the farmer will not experience dissonance, since he did not foresee the consequences of his choice and hence will not feel responsible. According to Cooper the crucial issue thus relates to the role of "foreseeability" in the production of dissonance. A previously reported experiment by Brehm (1959) had seemingly indicated that unforeseen consequences, or a *fait accompli*, could create dissonance. Cooper argues for a different interpretation of the results.

In 1970 the issue was attacked empirically by Brehm and Jones in the context of a free-choice situation. This experiment demonstrated the important role of foreseen as opposed to surprise consequences in the production of dissonance. Subsequent similar research by Cooper (1971) and Worchel and Brand (1972) provided further evidence for the importance of responsibility. In the context of forced-compliance situations Collins and Hoyt (1972) have also argued for the importance of responsibility in the production of dissonance.

If we return to the cognitive band, the self (+) is responsible for (?) advocating that the dull task is enjoyable (-), it is

evident from the perspective of balance theory that the subject will not wish to accept responsibility for the negative consequences. If he or she does so the cognitive linkage becomes positive and, according to the multiplicative rule (Cartwright & Harary, 1956), the band is imbalanced. We assume that two general factors will have a bearing on this matter. First are the polarities of the two objects of affective significance. The more positive the evaluation of the self and the more negative the evaluation of consequences the less willing the subject will be to accept responsibility. Although the language is different, this assumption is equivalent to Osgood and Tannenbaum's (1955) assumption that incredulity regarding assertions increases with an increase in the polarities of the associated objects of judgment. Second, we will assume (in agreement with Cooper, Note 1) that the more ambiguous the evidence regarding responsibility the less willing the subject will be to accept responsibility for negative consequences. What constitutes evidence for responsibility? Quite possibly a large number of circumstances—some of which are discussed later. We believe, however, that many of these circumstances relate to the degree of positiveness in the cognitive relation between the self and the consequences of the choice or act.

As Insko and Schopler (1967) have pointed out, a cognitive relation is a perceived relation between attitude objects or objects of affective significance. These relations are grossly categorized as positive, negative, or null. Positive relations involve a perceptual tendency to work toward unity of the associated attitude objects, and negative relations involve a perceptual tendency to work toward disunity of the associated objects. Thus "likes," "facilitates," "follows from," "stands by," and "is responsible for" are positive cognitive relations, and "hinders," "dislikes," and "contradicts" are negative cognitive relations. A null cognitive relation is simply the negation of a positive cognitive relation, for example, "is not responsible for." Insko and Schopler further make the obvious point that cognitive relations un-

doubtedly differ in the degree of positiveness or negativeness. Thus the stronger the tendency to perceive unity or sameness between two associated attitude objects the stronger the positive cognitive relation.

In terms of the present situation, we argue that it is important to consider the degree of positiveness in the cognitive relation between the self and the consequences of the act. One way to decrease the positiveness of this cognitive relation is to increase the time and physical distance between the self and the consequences of the act. Thus it is commonly observed that push-button warfare, although far more deadly than hand-to-hand combat, is, in fact, more impersonal. With hand-to-hand combat the positive cognitive relation between the self and the consequences of aggression is stronger and the evidence for responsibility thus more compelling. This observation is consistent with Milgram's (1965) finding that the maximum shocks declined with an increase in the "immediacy of the victim." The strength of the cognitive relation should theoretically also be affected by such things as the presence or absence of assumed future contact, the degree of similarity between self and victim, and the extent to which the victim is liked. There is, of course, no logical reason why assumed responsibility should be affected by the general positiveness of the cognitive relation, but we believe the generalization is nonetheless true. As will be indicated later, responsibility is an ill-defined and vague cognitive relation. Thus the possibility of confusion with other cognitive relations is fairly high.

In any event, the important point is that a remarkable number of circumstances may facilitate an individual's denial of responsibility for negative consequences. Cooper (Note 1) discusses the fact that numerous psychologists, philosophers, and writers have emphasized the general tendency to avoid assuming responsibility for "actions which lead to frustrating, uncomfortable or generally negative consequences" (p. 4). He points out that Horney (1950) considers the denial of responsibility an important aspect of neurotic personality development

and that Fromm (1941) makes the same point regarding normal personality development. Fromm's notion of "escaping from freedom" directly involves a denial of responsibility. Cooper also points out that existential thinkers such as May (1958) and Sartre (1956) come to a similar conclusion. According to Sartre, "most of the time, we flee anguish in Bad Faith" (p. 95). It also turns out that Shakespeare had the same idea:

"Fools by heavenly compulsion; knaves . . . by spherical predominance. . . . An admirable evasion of whoremaster man, to lay his goatish disposition to the charge of a star!" (*King Lear*, I, ii)

It indeed appears that most people in our society are reluctant to accept responsibility for negative consequences. From our perspective, however, this reluctance is intriguing because it follows from balance theory. Assuming a positive self and negative consequences, acceptance of responsibility would create imbalance.

Suppose, however, that circumstances, as yet unspecified, force an individual to accept responsibility. As previously noted, in such a situation the cognitive band is imbalanced. Assuming the evidence for responsibility is sufficiently compelling, balance may be restored by ceasing to regard the self as positive or ceasing to regard the counterattitudinal act as negative. Since lowering one's self-esteem would require a rather massive reorganization of the self-concept and associated attitudes structure, we regard a change in evaluation of the counterattitudinal act as more likely. This assumption is similar to Abelson and Rosenberg's (1958) principle of least effort in the restoring of consistency in a conceptual arena or attitude structure. How can the subject change his or her evaluation of the counterattitudinal act? Possibly the subject could decide that it is not too bad to mislead someone. Such a conclusion, however, is not very plausible. More plausibly, the subject could decide that "I did not mislead the other subject because I really did enjoy the task." Such a change would eliminate imbalance in the responsibility band involving the self. Also a further cognitive band is implicitly in-

volved: "My advocacy (+) pointed out that (+) the task is enjoyable (+)."

From a balance theory perspective, however, imbalance-resolving change in the enjoyableness of the task is dependent upon the subject's acceptance of responsibility for the counterattitudinal act. But what do we mean by the term "responsibility," or, more exactly, how do subjects in the Festinger and Carlsmith (1959) situation conceive of responsibility? At this point, it may be useful to review briefly Heider's (1958) five "levels" of responsibility. Heider's levels differ both in the restrictiveness of the implied definition of responsibility and in the extent to which responsibility is attributed to the environment rather than to the person. Fishbein and Ajzen (1973) label and describe these levels as follows:

- (1) *Association*: At the first and most primitive level, the actor is held responsible for all effects that are in any way associated with him.
- (2) *Commission*: At the next level he is held responsible if he was instrumental in producing the observed effects (even if he could not have foreseen them).
- (3) *Foreseeability*: At this level the actor is held responsible only if he could have foreseen the effects even though he might not have intended to produce them.
- (4) *Intentionality*: At the fourth level he is held responsible for effects he foresaw and intended.
- (5) *Justification*: Finally, at the fifth level he is held responsible only to the extent that his intended behavior was not justified, i.e., not caused by environmental factors beyond his control. (pp. 149-150)

On the fifth and most restrictive level, responsibility is more likely to be attributed to the environment than to the person. At this level Heider has what might be characterized as a "variance accounted for" definition of responsibility. Given certain environmental circumstances, if almost everyone acted in the same way, the individual in question is not responsible. According to Heider, it is at this level that "the criminal may blame the environment for his ill-fated career and thereby excuse himself" (1958, p. 114).

It strikes us as quite likely that the typical subject in the Festinger and Carlsmith (1959) situation initially conceives of responsibility as a confused mixture of the various levels, that is, as a fairly general positive cognitive relation. However, since

the first three levels will not enable the subject to deny responsibility and avoid imbalance, it is probable that one or both of the last two levels are quickly arrived at. Whether or not there is any kind of sequential shift, we will assume that the typical subject mainly conceives of responsibility in terms of a mixture of intentionality and variance accounted for, that is, Levels 4 and 5.

We are not in a position to ask what causes someone to accept responsibility for negative consequences. Theoretically, high choice provides evidence for some intentionality and also indicates that everyone might not have acted in the same way. Conversely, low choice provides evidence against intentionality and also indicates that everyone might have conformed to the pressure and acted in the same way. Therefore, choice conveys evidence for responsibility, regardless of whether responsibility is conceived of in terms of intentionality or variance accounted for. Is choice enough, however? We are inclined to think not. Except in free-choice situations where the subject chooses between desirable alternatives, we believe that choice per se provides rather weak cognitive glue. Although this assertion is somewhat at variance with the emphasis traditionally placed upon choice, it is interesting to note that Calder et al. (1973) did not obtain a main effect for choice. In neither the high-consequences nor the low-consequences conditions was there a general tendency for high choice to differ from low. Choice in combination with some additional circumstance or circumstances may, however, force acceptance of responsibility for negative consequences. These additional circumstances may include the expenditure of high effort preparing for or creating the negative consequences or the toleration of threatened or actual punishment in order to create the negative consequences. The results of Calder et al. indicate that the additional circumstance in the Festinger and Carlsmith (1959) situation is reward.

Reward has three functions in the Festinger and Carlsmith situation. First, reward provides evidence regarding the

extent to which other subjects would comply with the request and argue that the dull task is enjoyable. When the reward is large, it is plausible for the subject to assume that other subjects would have behaved in the same way. Therefore, from the perspective of the variance-accounted-for definition of responsibility, a large reward provides evidence for the denial of responsibility. Also, when the reward is small, it is not plausible for the subject to assume that other subjects would have behaved in the same way. A small reward thus provides evidence for the acceptance of responsibility. This may be true even in the case in which the consequences are partially unforeseen, that is, when there is only partial intentionality. As long as some of the consequences are foreseen, the occurrence of further unforeseen consequences may not, given the free choice of a small reward, deter the acceptance of responsibility.

The second function of reward is as evidence of the experimenter's concern with the experiment and thus his desire or intention that the waiting subject be misled. We believe that subjects are less likely to use a variance-accounted-for definition of responsibility than a definition in terms of one of the other four levels when they focus on the experimenter. This is partially because they are not motivated to be as analytic regarding the experimenter as they are regarding themselves, and partially because they have no evidence regarding how other experimenters would act in the same situation. Thus, when the reward is large it may be quite plausible to point a finger at the experimenter and indirectly absolve the self of responsibility. On the other hand, when the reward is low and insufficient such a shift in responsibility is much less compelling. Indeed the very minimal nature of the reward is some evidence that the experimenter did not exert undue pressure on the subjects and is not overly concerned with the experiment. Thus the low reward makes it less plausible to believe that "the experimenter is responsible."

The third function of the reward in the Festinger and Carlsmith (1959) paradigm

is as a *cue* to the enjoyableness of the task. If the subject agreed to argue that the task was enjoyable in return for only a low or minimal reward, the reward becomes a possible cue that the task was in fact enjoyed. Otherwise, how is the compliant behavior to be explained? A high reward, of course, carries no such cue value. The cue implications of low reward in the high-dissonance (high-choice, high-consequences) cell are particularly welcome since acceptance of the implications allows the subject to resolve imbalance in the cognitive band.

The above line of argument is in some respects reminiscent of Bem's (1967) attribution interpretation of dissonance effects. We, however, do not endorse his somewhat extreme position regarding the equivalence of actor and observer attribution and the unimportance of prior attitudes. Some recent research by Ross and Shulman (1973), in fact, calls into question Bem's denial of the importance of prior attitudes (cf. Bem & McConnell, 1970). On the other hand, if attribution theory is interpreted to mean that the subject goes through an inference process, perhaps even an "unconscious inference" process (Helmholtz, 1866/1925), then what we have said is certainly compatible with at least some aspects of attribution theory. The fact that observers in the Calder et al. (1973) study were able to track the triple interaction pattern certainly does not rule out the role of inferential processes. Actually, since both balance and attribution notions were put forth by Heider (1958), they are hardly strange bedfellows.

In any event, we are able to account for the attitude change that occurs in the high-dissonance, or "Oh my God what have I done," cell of the Calder et al. (1973) study. In the remaining three cells of Table 1, either choice was low or reward was high, or both. Thus in these cells the subject theoretically should have been more prone to resolve imbalance by denying responsibility for the negative consequences. It should be further noted that only the high-dissonance cell differed from control.

One further complicating factor relating to the high-dissonance cell involves the tension or nervousness that subjects may

have felt as a result of having prevented the waiting subject from studying for the quiz. Although the quiz was unforeseeable, information regarding it did make the consequences even more negative. On the other hand, the information regarding the quiz may have also produced what Brock (1963) calls "irrelevant tension." Brock found that "spreading of the alternatives" in a free-choice situation was more pronounced among children who had cheated in the immediately preceding situation and thus were more nervous and upset. There is therefore some reason to suppose that tension and nervousness act as a catalyst facilitating inconsistency or imbalance-resolving changes. If this is true the magnitude of the results in the Calder et al. (1973) high-dissonance cell may have been partially a function of such "irrelevant tension."

By way of restatement, we can regard the subjects in the high-dissonance cell as going through an inference process. While such a process may involve any of a number of steps or orders, one possibility is as follows. First, the subject confronts the negativity of the consequences and counter-attitudinal act. Second, since acceptance of responsibility would produce imbalance, there may be initial consideration of responsibility denial. Third, in view of the strong evidence for responsibility flowing from the joint existence of high choice and low reward, responsibility is tentatively accepted, thereby creating a potential or actual dilemma (imbalance). Fourth and finally, imbalance is avoided or reduced by the subject's deciding that the task is enjoyable and that the waiting subject was not misled. Other sequences are, of course, possible, but we nonetheless believe that the subject does go through an inferential process of some sort. Quite possibly the process occurs with considerable rapidity.

To this point we have offered a theoretical explanation for everything that is explained by dissonance theory in the Calder et al. (1973) experiment. It is only by virtue of its absence that dissonance theory is relevant to the other three cells of the design. Dissonance theory, for example, does not attempt to account for reinforcement

effects. It is true that none of the three remaining cells of Table 1, including the two reinforcement-effect cells, differs from control. Nonetheless we feel that they deserve explanation. The two reinforcement-effect cells differ significantly from each other, and, as previously indicated, other investigators have found reinforcement effects. What follows, therefore, is an attempt to account for the additional between-cells variance in Table 1. This explanation is based on a postulated sequential-inference process.

We will begin by noting that Calder et al. (1973) only obtained a reinforcement effect in the high- or negative-consequences condition. In the low-consequences condition the low-choice means are actually reversed in direction, although the reversal is small and nonsignificant. Thus, contrary to common sense expectations, high or negative consequences are important in the prediction of Calder et al.'s reinforcement effect. This in turn suggests to us that the reinforcement effect is somehow related to the subject's denial of responsibility.

We consider first the low-reinforcement cell. Since choice was low, the evidence for the subject's assumption of responsibility is not very strong. Nonetheless, the situation is somewhat equivocal in view of the low reward. As we noted previously, low reward makes it difficult for the subject to shift responsibility to the experimenter and to believe that other subjects would have acted in the same way. Thus the subject may for an instant consider accepting responsibility. During this period the subject will be forced by the potential imbalance to consider the negativeness or awfulness of the counterattitudinal act. In view of the negative polarity of the counterattitudinal act, the partial unforeseeability of the consequences, and the lack of high choice, however, the subject gets out of the dilemma (imbalance) by denying responsibility. The subject may reason that other subjects would have been pressured into doing just what he or she did. Thus what we are postulating is a sequential-inference process, in which responsibility is first considered and then rejected. Further, we believe that such sequential

inference results in an even more negative evaluation of the counterattitudinal act.

There are two possible reasons why this is so. First, as a result of the subject's having considered acceptance of responsibility for the counterattitudinal act, the negativity of that act becomes particularly salient. Second, by assuming that his or her counterattitudinal act is extremely negative, the subject can account for the strong denial of responsibility. The subject, in effect, reasons that "since I am avoiding something it must be pretty awful." What we are postulating is a type of attribution based not on overt behavior, but on a newly formed belief or cognitive linkage ("I am not responsible" or "It's not my fault"). If the subject regards the counterattitudinal act as highly negative, then the fact that the counterattitudinal act is positively associated with the task indicates, according to the multiplicative rule, that the task should be regarded negatively. This then accounts for the low mean of 9.60 obtained in the low-reinforcement cell.

The situation in the high-reinforcement cell is much less threatening for the subject. Since there is both low choice and high reward, there is even less reason for the assumption of responsibility. The subject thus may not be severely threatened with the possibility of personal responsibility for the negative consequences. As a result, the negativity of those consequences may not be salient, and the subject's denial of responsibility may be less forceful. A less vehement denial of responsibility would in turn have less of the attribution implication that "since I am avoiding something it must be pretty awful."

The same argument made for the low-reinforcement cell could also be made for the low-dissonance cell. In the latter instance, however, the threat of responsibility would come from high choice.

To recapitulate briefly, we have argued that the Festinger and Carlsmith (1959) paradigm can best be viewed from the perspective of a cognitive band: The self (+) is responsible for (?) advocating that the dull task is enjoyable (?). Under high or negative consequences the counteratti-

tudinal advocacy becomes negative. In the high-dissonance cell the combination of high choice and low reward forces the subject to consider accepting responsibility, thus creating a positive cognitive linkage and potential or actual imbalance. The low reward also acts as a cue, indicating that the task was enjoyable. Imbalance is resolved or avoided by a change in the evaluation or the task, thereby allowing the subject to decide that the other subject was not misled and that the counterattitudinal advocacy was not negative. In the low-reinforcement cell, where there is both low choice and low reward, the threat of responsibility suggested by the low reward leads the subjects first to consider the negativity of the counterattitudinal advocacy and then to forcefully deny responsibility. The heightened salience of the negative counterattitudinal advocacy (or high consequences) plus the attributional explanation of the forceful denial of responsibility leads, via a balancing generalization of affect, to a more negative evaluation of the task.

For now, in the hope of not straining the reader's credulity to the breaking point, we will restrain ourselves from further speculation concerning the role of the affective value of high reward. Such speculation may appear more appropriate after presentation of the data from the present experiment.

The main purpose of the experiment is to provide an empirical test of the balance interpretation of the Calder et al. (1973) results. Since this interpretation is presented in terms of rapidly occurring cognitive processing, a definitive and direct test is difficult to achieve. The present test is less direct. What we have done is to search for a variable which, although operationally distinct from reward, has roughly similar implications with regard to the subject's assumption of responsibility and similar cue properties with regard to enjoyment of the task. We believe that such a variable is the investment of time, which was suggested to us by Wicklund, Cooper, and Linder (1967).

In the Wicklund et al. experiment subjects anticipated waiting either 1 minute or

15 minutes for exposure to a counterattitudinal message. Waiting 15 minutes required that the subjects remain longer than the 30-minute period to which they had initially agreed. All subjects were tested under high-choice conditions, and there was no manipulation of consequences. None of the subjects actually waited for the exposure; the manipulation was of anticipated waiting. The results indicated significantly more agreement with the anticipated message in the 15-minute than the 1-minute condition. Wicklund et al. considered the anticipated-waiting manipulation to be one of anticipated effort. Certainly, however, time relates to effort only in a metaphorical sense. But beyond this, the crucial aspect of the Wicklund et al. manipulation may relate to the fact that the 15-minute subjects were asked to wait longer than they had originally agreed to wait. They were thus asked to invest more time. For that reason we would prefer to regard the manipulation as an investment of time or an anticipated investment of time. Admittedly, however, even this terminology does not reflect the fact that there was a violation of expectations for the 15-minute subjects.

It was argued previously that reward has three functions in the Festinger and Carlsmith (1959) paradigm. First, reward provides evidence regarding the extent to which other subjects would comply with the request and argue that the dull task is enjoyable. Second, reward provides evidence regarding the experimenter's concern with the experiment and thus his desire that the waiting subject be misled. Third, reward is a cue to the enjoyableness of the task. We believe that the first and third of these functions are also true for investment of time. When the investment of time is high, it is not plausible for the subject to assume that other subjects would have behaved as he or she did. High investment of time thus provides evidence for the acceptance of responsibility. Also, if the subject agrees to argue that the task is enjoyable even when high investment of time is required, the investment of time becomes a cue that the task was in fact enjoyed. Otherwise, how is the compliant

behavior to be explained? We therefore predict the same triple interaction that was obtained by Calder et al. (1973). The "dissonance" effect for investment to time should occur when there is both high choice and negative consequences. Note that this prediction is, superficially at least, in disagreement with the results of Wicklund et al. (1967), since they had no manipulation of consequences.

Beyond this, however, investment of time has one rather salient difference from reward. Low investment of time, unlike high reward, does not have any affective value. Thus any difference between the present data and that of Calder et al. will provide some circumstantial information regarding this matter.

METHOD

Subjects

The subjects were 126 female undergraduates enrolled in the introductory psychology course at the University of North Carolina who had signed up for a "Measures of Performance" experiment. Data from 12 of these subjects were not included in the analyses. One subject (low choice, low consequences, high time) did not attempt to convince the stooge of the task's enjoyableness. Three subjects (all high choice, high time) refused to talk to the waiting subject. Finally, debriefing revealed that eight subjects had suspicions regarding some aspect of the procedures. These subjects were scattered across seven of the eight cells of the design. The experimenter's informal impression was that suspicion was to a large extent a function of having participated in previous experiments. There was, for example, almost no suspicion from subjects tested early in the semester.

Independent Variables

The three independent variables (choice, consequences, and investment of time) were manipulated in the experiment. Choice was manipulated by either giving the subject a choice or not giving the subject a choice as to whether or not she would tell the waiting subject that the task was enjoyable. The consequences variable was manipulated in the same way as Calder et al. (1973). In the high-consequences condition the male stooge was initially skeptical and then became convinced that the experiment was interesting and enjoyable. He also stated that since the experiment was enjoyable he would stay rather than leave to study for a quiz, as he had initially intended. In the low-consequences condition the stooge remained skeptical and unconvinced. The investment of time manipulation was modeled after the Wicklund et al. (1967)

manipulation—the main difference being that our subjects actually did wait differing amounts of time rather than just anticipating a given wait. In the high-time condition the subjects had to wait 15 minutes before they could meet the next subject, and this extra wait meant that they would not be through until ten minutes after the end of the time period for which they originally signed up. In the low-time condition the subjects only had to wait 1 minute while the experimenter checked to see if the next subject had arrived. Subjects in this condition expected to be finished within the time for which they had originally signed up.

Beyond the eight groups or conditions created by the above factors, three additional groups were tested. One of these was a standard control group. After working on the dull task, subjects simply filled out the final questionnaire. There was no manipulation of choice, consequences, or investment of time. The other two groups were tested under high-time-anticipation, high-choice, no-consequences conditions or low-time-anticipation, high-choice, no-consequences conditions. After working on the task, subjects were given a choice as to whether or not they would wait (either 1 or 15 minutes) and talk to the other subject. They were then given the final questionnaire. There was no manipulation of consequences. The intent was to see whether the anticipation of high or low investment of time, high choice, and an unmanipulated anticipation of consequences would produce results similar to those of Wicklund et al. (1967).

All of the above groups, except for the control, contained 9 subjects. The control group contained 24.

Procedure

Since the procedure so closely paralleled the procedure and script of the Calder et al. (1973) study, a detailed description will not be repeated here. Departure from the Calder et al. procedure mainly occurred for the investment of time (as opposed to reward) manipulation. In the high-time condition the subjects were told,

I think we have one male subject signed up at — (the end of the half-hour period for which the subject had signed up; e.g., 12:30 for those who had started at 12:00), but I've got some data to work on and won't be able to start until — (5 minutes beyond the end of credited time; e.g., 12:35). That means that you would be through by about — (10 minutes after the allotted time; e.g., 12:40).

In the low-time condition the subjects were told,

I think we have one male subject signed up to start right about now, so you could go in to talk to him as soon as I make sure he's shown up. That means you would be through by about — (5 minutes before the full period for which the subject had contracted; e.g., 12:25 for those starting at 12:00).

Assessment of Dependent Variables

After retrieving the subject from the waiting room, the experimenter set the stage for assessment of the dependent variables:

There is one more thing that I would like you to do. The Psychology Department is taking a survey of students' reactions to psychology experiments. The Department has asked me to ask you if you would mind coming by the office for just a moment. I think they have some sort of brief questionnaire they want you to fill out.

At the office, the subject was greeted by a secretary who administered the questionnaire. Subjects were met by the experimenter as they came out of the office and fully debriefed.

The first page of the "Psychology Department Questionnaire" contained the following statement:

As part of a program to evaluate the effectiveness of the requirements for outside participation in psychology courses, the Psychology Department is conducting a survey of student reactions to a random sample of experimental sessions. Your answers to the following questions will help give us some idea of the way students typically react to experiments. Please answer each question honestly, giving your own opinion, and check to make sure you complete the entire questionnaire. When you have finished and checked over your answers, deposit this questionnaire in the waiting room in the box marked "Psychology Department Questionnaires." All responses are anonymous, so *do not sign your name*.
Thank you for your cooperation.

The second page of the questionnaire contained a single 31-point rating scale preceded by the question, "How enjoyable was the experimental task you performed?" The low end of the scale was labeled "extremely unenjoyable" and the high end "extremely enjoyable." (The data reported in Table 1 from the Calder et al. (1973) study were based on ratings on a 36-point, rather than a 31-point scale.) Subsequent pages of the questionnaire contained various filler items, such as amount learned in the experiment.¹

The final page of the questionnaire was designed to assess the manipulations and possible misgivings. The page began with the question, "Did you perform any other duties (such as assisting the experimenter in some way) in addition to the main experimental task? () Yes () No." If the subject answered "Yes" she was instructed to answer the questions below. These included, "How much freedom did you have to choose whether or not you would perform this additional function?" "How successful were you in performing this additional function?" "To what extent did you have misgivings about performing this function?" Answers

¹ An analysis of variance of this "cover item" revealed the same significant triple interaction that was found for the main dependent variable.

TABLE 2
MEAN REPORTED ENJOYMENT

Investment of time	High consequences		Low consequences	
	High choice	Low choice	High choice	Low choice
15 minutes	18.67	10.67	15.67	15.33
1 minute	8.44	13.78	16.33	12.67

Note. Control = 14.12, high time-anticipation = 15.25, low time-anticipation = 15.75. Ratings were made on a 31-point scale in which higher numbers indicate more enjoyment.

were made on 7-point scales in which low numbers indicated the least amount of the attribute in question (freedom, success, misgivings).

RESULTS

Perceived Choice and Consequences

A least squares analysis of variance of the perceived-choice ratings yielded only a main effect for choice, $F(1, 59) = 102.94$, $p < .01$. Ratings on the 7-point scale indicated greater choice or freedom in the high-choice condition ($\bar{X} = 6.75$) than in the low-choice condition ($\bar{X} = 2.61$).

A least squares analysis of variance of success in carrying out the "additional function" resulted in a significant main effect for consequences, $F(1, 59) = 41.02$, $p < .01$. There was more reported success in the high-consequences condition ($\bar{X} = 5.99$) than in the low-consequences condition ($\bar{X} = 3.98$). This analysis resulted in one more significant effect, a main effect for choice, $F(1, 59) = 9.63$, $p < .01$. High-choice subjects reported more success ($\bar{X} = 5.48$) than did low choice subjects ($\bar{X} = 4.50$). The stooge's rating of the adequacy of the counterattitudinal performance, however, did not agree with the subjects' perceptions—the choice F being less than 1.

Enjoyment of the Task

Mean reported enjoyment of the task is contained in Table 2. A least squares analysis of variance resulted in only one significant effect, the triple interaction, $F(1, 64) = 6.04$, $p < .017$ (this triple interaction was significant in a multivariate

TABLE 3
MEAN REPORTED MISGIVINGS

Investment of time	High consequences		Low consequences	
	High choice	Low choice	High choice	Low choice
15 minutes	2.44	2.25	3.44	2.00
1 minute	4.29	2.38	2.12	2.56

Note. Ratings were made on a 7-point scale in which high scores indicate greater misgivings.

analysis of variance covering all of the assessments). Breakdown of this Choice $\times$ Consequences $\times$ Investment of Time interaction resulted in a significant Choice $\times$ Investment of Time interaction within high consequences, $F(1, 32) = 8.76, p < .01$, that was nonsignificantly reversed within low consequences, $F(1, 32) < 1$. The double interaction within high consequences resulted from a significant dissonance effect with high choice, $t(64) = 3.02, p < .01$, and a nonsignificant reinforcement effect with low choice ($t < 1$).

The control mean, 14.12, differs significantly both from the high-dissonance mean, $t(87) = 2.31, p < .05$, and the low-dissonance mean, $t(87) = 2.4, p < .05$.

After completion of the experiment it was suggested to us that perhaps a more accurate control procedure would involve telling subjects that they were, in fact, control subjects and that there was potentially more to the experiment than just working on the task. In any event, we ran two additional control groups, one of which simply worked on the task (as above), and another of which was given the additional information. Each of these control groups contained 12 subjects. Mean enjoyment was identical (13.91) for the two groups.

The means for the high-time-anticipation condition (15.25) and low-time-anticipation condition (15.75) do not differ significantly from each other or from control.

Reported Misgivings

Mean reported misgivings are reported in Table 3. High numbers indicate greater misgivings. A least squares analysis of variance resulted only in a significant

triple interaction, $F(1, 59) = 4.34, p < .05$. This triple interaction to a large extent results from the pattern within the dissonance cells (high choice, high consequences). Low dissonance resulted in the most reported misgivings and high dissonance in low reported misgivings.

The Calder et al. (1973) study found that observers attributed high misgivings in the high-dissonance cell, while actors reported low misgivings in the high-dissonance and high-reinforcement cells. The present data are comparable, in that high dissonance (relative to low dissonance) resulted in low reported misgivings. These data, however, do not indicate a similar trend within the reinforcement cells—possibly because there was no reinforcement effect (unlike the Calder et al. study).

DISCUSSION

The results of the experiment substantiate our major expectation of triple interaction for rated enjoyableness of the task. With high consequences there was a significant double interaction between choice and investment of time that was nonsignificantly reversed with low consequences. The significant double interaction resulted from a significant dissonance effect with high choice and a nonsignificant reinforcement effect with low choice. Both the high- and low-dissonance cells differed significantly from control.

The high-time-anticipation and low-time-anticipation conditions did not differ significantly from each other. Apparently manipulated anticipated time, together with manipulated anticipated consequences and high choice, are not sufficient to produce the dissonance effect, as suggested by the Wicklund et al. (1967) experiment.

We are, of course, gratified that investment of time and reward both interact with choice and consequences in the context of the Festinger and Carlsmith (1959) paradigm. There are, however, some interesting differences between the present results and the Calder et al. (1973) results—results that have a bearing upon the previously alluded to affective value of high reward. In the present experiment the control mean is

intermediate to the high- and low-dissonance means, and differs significantly from both. In the Calder et al. study, however, the control mean differs only from the high-dissonance mean. The low-dissonance mean is actually slightly above the control mean (see Table 1). This fact is potentially embarrassing to our previous argument that the threat of responsibility, which occurs when there is low choice and low reward (low-reinforcement cell) and high choice and high reward (low-dissonance cell), is followed by a forceful denial of responsibility and subsequent devaluation of the task. Fortunately for that argument, the low-dissonance cell of the present data is indeed below control. But why did this not occur in the Calder et al. experiment? A possible answer relates to the affective value of the high reward. In the context of high consequences, the high reward may have taken on added salience due to its implication that the experimenter carried responsibility for what happened. This in turn might counteract the aforementioned tendency to downgrade the task as a result of responsibility threat and denial.

We realize that this argument is highly speculative. The argument, however, does allow us to account not only for the low mean in the low-dissonance cell but also for the lack of a reinforcement effect. The reinforcement effect in the Calder et al. (1973) data is significant; in the present data it is not. Furthermore, this lack of a reinforcement effect apparently is due to the fact that the mean in the high-reinforcement cell is not above control. In the Calder et al. study an increase above control would have been facilitated by the affective value of the high reward. In the present study there was no reward.

In summary, we have attempted to account for the high consequences data of the present experiment and of the Calder et al. experiment with a three-part argument. First, we have an argument for the high-dissonance cell. According to this argument, the subject decides that the task is enjoyable in order to be able to accept responsibility for the counterattitudinal act,

as the high choice and low reward indicate that she should.

Second, we have an argument directly relating to the low-dissonance and low-reinforcement cells, and indirectly relating to the high-reinforcement cell. In the low-dissonance and low-reinforcement cells there is conflicting evidence regarding responsibility that leads first to a consideration of acceptance of responsibility and then a forceful denial of responsibility. This sequence increases the salience of the negative consequences and suggests, via attributional inference, that the avoided consequences are indeed negative. Thus the rating of the task is lowered in the low-reinforcement and low-dissonance cells. This argument is indirectly related to the high-reinforcement cell, where the lack of evidence for responsibility does not result in such a forceful denial of responsibility and thus a lowered evaluation of the task.

Third and finally, we have an argument regarding the affective value of the high reward that was present in the Calder et al. (1973) experiment and is absent in the current experiment. The affective value of the high reward prevented the low dissonance cell of the Calder et al. study from going below control (as it did in the present study) and produced a higher rating in the high-reinforcement cell and thus a significant reinforcement effect (unlike the present study).

Toward the beginning of this article we made the rather ambitious statement that balance theory could account for all dissonance effects. So far we have only discussed forced-compliance situations, in particular the role of reward, choice, and investment of time in forced-compliance situations. It should be apparent that a similar analysis could be made concerning the role of punishment and effort in forced-compliance situations. The salient issue here revolves around the extent to which punishment and effort inherently involve negative consequences. Our guess is that punishment does involve negative consequences and that effort does not. But what about free-choice situations? A balance interpretation of free-choice situations is very

straightforward. Two cognitive bands are involved: the self (+) selects or chooses (+) alternative *A* (?), and the self (+) rejects or does not choose (-) alternative *B* (?). According to the multiplicative rule, alternative *A* should move in the positive direction and alternative *B* should move in the negative direction; that is, "spreading of the alternatives" should occur. It should also be noted that an implicit choice could produce some spreading before the explicit choice, but the spreading will be most marked after the actual choice, when the linking cognition is strongest. Thus balance theory and dissonance theory agree in regarding the spreading of the alternatives as a post-decisional, rather than a predecisional, effect. It is also true that this theoretical account, like dissonance theory, is potentially embarrassed by regret (Walster, 1964).

In general, we are impressed by balance theory's ability to account not only for dissonance effects but also for such diverse additional effects as attitudinal verbal reinforcement (Insko & Cialdini, 1971), affective conditioning (Insko & Schopler, 1972), interpersonal attraction (Aronson & Cope, 1968), and overevaluation of in-group products (Ferguson & Kelley, 1964). Despite its lack of general popularity or acceptance, balance theory remains one of the most generally applicable theories in all of social psychology. The main problem with the theory is that it has not been developed to specify which cognitive bands are relevant to which situations. There is nothing in balance theory per se that tells us, for example, that a cognitive band involving self-responsibility for negative consequences is relevant to the Festinger and Carlsmith (1959) situation. Tests of balance theory, and indeed of all theories, follow an "if, then" format. If certain cognitive bands exist, then certain testable predictions follow. The problem with balance theory is that we are frequently very ignorant regarding the "if" conditions.

The above perspective on balance theory is exactly analogous to the generally accepted perspective on various hedonistic

theories (exchange theory, for example). Exchange theory is a general position regarding the importance of rewards and costs in the determination of behavior. The theory, however, does not say what individuals regard as rewards and costs in particular situations. Thus what appears to be an empirical test of exchange theory is in reality an empirical test of the assumption that certain rewards and costs are operating in a certain situation. Similarly, what appears to be a test of balance theory is in reality a test of the assumption that certain cognitive bands are operating in a given situation. Thus we theorized that a band suggesting possible self-responsibility for counterattitudinal advocacy and associated negative consequences is important in the Festinger and Carlsmith (1959) paradigm. This in turn led us to expect that, contrary to the Wicklund et al. results, investment of time should interact with choice and consequences in a fashion roughly analogous to reward. This, in fact, happened. We would be the last to claim that such positive results confirm a point of view. The last word on what has historically proven to be both a puzzling and intriguing phenomenon has undoubtedly not been said.

REFERENCE NOTE

1. Cooper, J. *Personal responsibility and cognitive dissonance*. Unpublished manuscript, 1968.

REFERENCES

- Abelson, R., & Rosenberg, M. Symbolic psychology: A model of attitudinal cognition. *Behavioral Science*, 1958, 3, 1-13.
- Aronson, E. Dissonance theory: Progress and problems. In R. Abelson, E. Aronson, W. McGuire, T. Newcomb, M. Rosenberg, & P. Tannenbaum (Eds.), *Theories of cognitive consistency: A sourcebook*. Chicago: Rand McNally, 1968.
- Aronson, E., & Cope, V. My enemy's enemy is my friend. *Journal of Personality and Social Psychology*, 1968, 8, 8-12.
- Bem, D. J. Self-perception: An alternative interpretation of cognitive dissonance phenomena. *Psychological Review*, 1967, 74, 182-200.
- Bem, D. J., & McConnell, H. K. Testing the self-perception explanation of dissonance phenomena: On the salience of premanipulation attitudes. *Journal of Personality and Social Psychology*, 1970, 14, 23-31.
- Brehm, J. W. Increasing cognitive dissonance by a *fait accompli*. *Journal of Abnormal and Social Psychology*, 1959, 58, 379-382.

- Brehm, J. W., & Cohen, A. R. *Explorations in cognitive dissonance*. New York: Wiley, 1962.
- Brehm, J. W., & Jones, R. A. The effect on dissonance of surprise consequences. *Journal of Experimental Social Psychology*, 1970, 6, 420-431.
- Brock, T. C. Effects of prior dishonesty on post-decision dissonance. *Journal of Abnormal and Social Psychology*, 1963, 66, 325-331.
- Brown, R. *Social psychology*. New York: Free Press, 1965.
- Calder, B. J., Ross, M., & Insko, C. A. Attitude change and attitude attribution: Effects of incentive, choice, and consequences. *Journal of Personality and Social Psychology*, 1973, 25, 84-99.
- Cartwright, D., & Harary, F. Structural balance: A generalization of Heider's theory. *Psychological Review*, 1956, 63, 277-293.
- Cohen, A. R. An experiment on small rewards for discrepant compliance and attitude change. In J. W. Brehm & A. R. Cohen (Eds.), *Explorations in cognitive dissonance*. New York: Wiley, 1972.
- Collins, B. E. Counterattitudinal behavior. In R. Abelson, E. Aronson, W. McGuire, T. Newcomb, M. Rosenberg, & P. Tannenbaum (Eds.), *Theories of cognitive consistency: A sourcebook*. Chicago: Rand McNally, 1968.
- Collins, B. E., & Hoyt, M. F. Personal responsibility-for-consequences: An integration and extension of the "forced-compliance" literature. *Journal of Experimental Social Psychology*, 1972, 6, 558-593.
- Cooper, J. Personal responsibility and dissonance: The role of foreseen consequences. *Journal of Personality and Social Psychology*, 1971, 18, 354-363.
- Cooper, J., & Worchel, S. Role of undesired consequences in arousing cognitive dissonance. *Journal of Personality and Social Psychology*, 1970, 16, 199-206.
- Elms, A. C., & Janis, I. L. Counter-norm attitudes introduced by consonant versus dissonant conditions of role-playing. *Journal of Experimental Research in Personality*, 1965, 1, 50-60.
- Ferguson, C. K., & Kelley, H. H. Significant factors in over-evaluation of own-group's product. *Journal of Abnormal and Social Psychology*, 1964, 69, 223-228.
- Festinger, L., & Carlsmith, J. Cognitive consequences of forced compliance. *Journal of Abnormal and Social Psychology*, 1959, 58, 203-210.
- Fishbein, M., & Ajzen, I. Attribution of responsibility: A theoretical note. *Journal of Experimental Social Psychology*, 1973, 9, 148-153.
- Fromm, E. *Escape from freedom*. New York: Harper, 1941.
- Heider, F. Attitudes and cognitive organization. *Journal of Psychology*, 1946, 21, 107-112.
- Heider, F. *The psychology of interpersonal relations*. New York: Wiley, 1958.
- Helmholtz, H. v. *Treatise on physiological optics* (Vol. 3) (J. P. C. Southall, Ed.). Ithaca, N. Y.: Optical Society of America, 1925. (Originally published, 1866.)
- Horney, K. *Neurosis and human growth: The struggle toward self-realization*. New York: Norton, 1950.
- Insko, C. A., & Cialdini, R. B. *Interpersonal influence in a controlled setting: The verbal reinforcement of attitude*. Morristown, N. J.: General Learning Press, 1971.
- Insko, C. A., & Schopler, J. Triadic consistency: A statement of affective-cognitive-conative consistency. *Psychological Review*, 1967, 74, 361-376.
- Insko, C. A., & Schopler, J. *Experimental social psychology: Text with illustrative readings*. New York: Academic Press, 1972.
- May, R. Contributions of existential psychotherapy. In R. May, E. Angel, & H. Ellenberger (Eds.), *Existence: A new dimension in psychiatry and psychology*. New York: Basic Books, 1958.
- Melson, W. H., Calder, B. J., & Insko, C. A. The social psychological status of reward. *Psychonomic Science*, 1969, 17, 240-242.
- Milgram, S. Some conditions of obedience and disobedience to authority. *Human Relations*, 1965, 18, 57-76.
- Nel, E., Helmreich, R., & Aronson, E. Opinion change in the advocate as a function of the persuasibility of his audience: A clarification of the meaning of dissonance. *Journal of Personality and Social Psychology*, 1969, 12, 117-124.
- Osgood, C. E., & Tannenbaum, P. H. The principle of congruity in the prediction of attitude change. *Psychological Review*, 1955, 62, 42-55.
- Rosenberg, M. J. Cognitive structure and attitudinal affect. *Journal of Abnormal and Social Psychology*, 1956, 53, 367-372.
- Rosenberg, M. J. When dissonance fails: On eliminating evaluation apprehension from attitude measurement. *Journal of Personality and Social Psychology*, 1965, 1, 28-42.
- Rosenberg, M. J., & Abelson, R. P. An analysis of cognitive balancing. In C. Hovland & M. Rosenberg (Eds.), *Attitude organization and change*. New Haven, Conn.: Yale University Press, 1960.
- Ross, M., & Shulman, R. F. Increasing the salience of initial attitudes: Dissonance versus self-perception theory. *Journal of Personality and Social Psychology*, 1973, 28, 138-144.
- Sartre, J. P. *Being and nothingness: An essay on phenomenological ontology* (H. Barnes, trans.). New York: Philosophical Library, 1956.
- Walster, E. The temporal sequence of post-decision processes. In L. Festinger (Ed.), *Conflict, decision and dissonance*. Stanford, Calif.: Stanford University Press, 1964.
- Wicklund, R. A., Cooper, J., & Linder, D. E. Effects of expected effort on attitude change prior to exposure. *Journal of Experimental Social Psychology*, 1967, 3, 416-428.
- Worchel, S., & Brand, J. Role of responsibility and violated expectancy in the arousal of dissonance. *Journal of Personality and Social Psychology*, 1972, 22, 87-97.

Perturbation Model for Letter Identification

George Wolford
Dartmouth College

A model is presented which provides a reasonably good quantitative account of the data from a series of tachistoscopic report experiments in which retinal locus, spacing, and array position were systematically varied. Potential extensions of the model to other paradigms are also suggested. The main assumptions of the model are that features are extracted from visual displays in unlimited parallel channels as a function of retinal locus; that a processor, working serially, attempts to identify letters from the extracted features; and that during processing, perturbations occur in the spatial ordering of the features.

The primary purpose of this article is to describe, justify, and elaborate a preliminary quantitative model for a subset of the letter-identification process. The model was formulated as a computer simulation and was specifically designed to account for the data from a series of tachistoscopic full-report experiments, in which retinal location, array position, and spacing were systematically varied. The article begins with a consideration of some of the recent theoretical advances in the area and a description of the empirical base of the model. The model is then described, along with an account of its correspondence to the empirical data. In the final sections it is suggested how the model can be applied to other paradigms, and some comments on its relationship to other models are offered.

MODELS OF VISUAL PROCESSING

A number of very promising visual information-processing models have appeared in the last few years. Estes (1972) presented an interactive-channels model to

This article was written while the author was visiting the Center for Human Information Processing at the University of California, San Diego. The center (under partial support from National Institute of Mental Health grant MH-15828) was generous in its assistance. Thanks are due William K. Estes, Samuel Hollingsworth, and David E. Rumelhart for their advice and assistance.

Requests for reprints should be sent to George Wolford, Department of Psychology, Dartmouth College, Hanover, New Hampshire 03755.

account for the data from a comprehensive detection experiment. The main assumptions of the model are (a) information is extracted from visual displays in parallel at the feature level, (b) the density of detectors for each feature type decreases with increasing distance from the center of the fovea, and (c) decisions are made on the basis of the features extracted at the various retinal locations. Many of the quantitative aspects of the model have not been fully specified, but a number of predictions can be deduced from the assumptions. The first is that only confusable items will interfere with one another. Their mutual interference arises over competition for the same set of feature detectors. Items that are not confusable would presumably not share any common features and would not, therefore, compete for the same set of feature detectors. In a detection experiment, then, the number of noise elements in the display would not be a relevant variable unless the noise elements were confusable with the targets. This prediction was borne out by Estes (1972). The second prediction is that the amount of interference among confusable items will increase as the items move into the periphery. A number of investigators have found evidence for increased letter interaction with increased distance from the center of the fovea (Estes & Wolford, 1971; Kolers & Rosner, 1960; Matthews, 1973). The interactive-channels model has

not yet been applied to other paradigms, such as full report or partial report.

Both Gardner (1973) and Shiffrin and Geisler (1973) developed versions of an independent parallel-channels model. The ensuing discussion focuses primarily on the version proposed by Shiffrin and Geisler, since it was developed in greater quantitative detail. In addition, their computer simulation bears certain similarities to the model that will be developed later. The basic assumptions of their model follow: (a) Each letter in a display has some independent probability of being recognized. (b) If the letter is not recognized, each feature of that letter has some independent probability of being detected. (c) A correct response is made if the entire target letter is recognized, or if one of the sets of detected features represents a subset of the target presented. Once again confusions represent the primary source of error in detection tasks, since confusable letters necessarily share some similar subsets. Nonconfusable noise letters have no effect, since they never lead to a subset of either target. With few parameters, Shiffrin and Geisler provided an excellent quantitative account of the data in Estes (1972). Their model has not yet been applied quantitatively to other paradigms.

The final model under consideration was formulated by Rumelhart (1970) and is referred to as a limited-capacity parallel model. The assumptions of the model are as follows: (a) Information from a visual display is stored in some iconic form but decays rapidly with time. (b) Information is extracted from the store, in the form of features, at a rate dependent on the amount of attention allocated to a particular location. (c) Observers have a finite amount of attention and can control its allocation. (d) The amount of information required per item to insure maximum performance varies with the paradigm. For instance, more information per item is required to respond correctly in a report task than in a detection task. Confusions are a prominent source of error, but nonconfusable items can also lead to errors, since they compete for the limited amount of avail-

able attention. Rumelhart's model provides an accurate quantitative account of performance (overall per cent correct) in a number of different paradigms.

While the differences among the models are testable experimentally (Estes, 1974; Gardner, 1973), they nevertheless share a number of primary assumptions. All three models assume that information is extracted in parallel and that the extraction process works at the level of features. In addition, they all assume that confusions are a major source of error in detection tasks (this is in part a corollary of the assumption of feature extraction). The success of these models in accounting for the data from various detection paradigms argues well for the feasibility of the shared assumptions. None of the models has been developed to make predictions about the specific nature of performance in report paradigms, for example, the shape of the serial position curves, or the effects of such variables as retinal location, array position, and spacing. The model developed in this article is primarily concerned with accounting for the effects of these variables and borrows heavily from the preceding models.

A number of theorists in the area have avoided report paradigms, partly because of their belief that those paradigms are contaminated by higher order processes, such as acoustic short-term memory (Gardner, 1973; Shiffrin & Geisler, 1973). Their view is that short-term memory capacity places an upper bound on the number of letters that can be reported in the full-report procedure, and further that the upper bound is less than the number of letters actually "seen." Since their interest has primarily been in studying what is actually "seen," report tasks are inappropriate. There is, however, little direct evidence that short-term memory is involved in report tasks. Wolford and Hollingsworth (1974a) argued against short-term memory restrictions, based on two considerations. The first reason was that the performance asymptote in the standard full-report procedure is approximately 4.0-4.5 letters (Estes & Taylor,

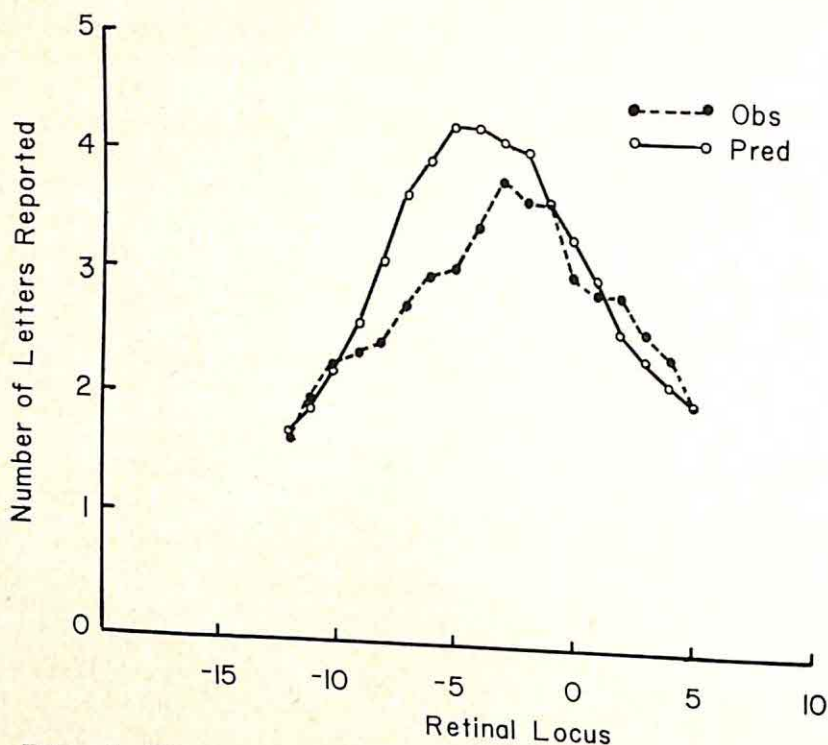


FIGURE 1. The mean number of letters reported as a function of the retinal placement of a display. (Abscissa numbers refer to elite type spaces, 5 spaces equal 1.05° . Abbreviations: Obs = observed; Pred = predicted.)

1964; Sperling, 1959), substantially below most estimates of acoustic short-term memory capacity, which run about 7.0 letters (Sperling & Speelman, 1970). The second reason was that incorrect responses from their full-report procedure yielded no evidence of acoustic confusions above a chance level, even though visual confusions abounded. With very few exceptions, most investigators have argued that short-term memory is normally acoustically based and have shown that acoustic confusions are prominent, particularly at short retention intervals, regardless of the mode of presentation (Conrad, 1964; Sperling & Speelman, 1970).

It is possible, of course, that the performance asymptote in full-report procedure is the result of a capacity limitation in a short-term memory with a nonacoustic representation. The existence of a visual short-term memory has been argued by Estes (1973), among others. The two preceding arguments were specific to acoustic short-term memory. One final argument against short-term memory capacity as a major influence in full-report

procedure is not specific to any particular type of representation in short-term memory.

Wolford and Hollingsworth (1974c) presented linear arrays of nine letters each at 18 different retinal locations. Figure 1 illustrates the mean number of letters reported per string as a function of the string's retinal locus. The numbers on the abscissa refer to the retinal locus of the leftmost letter of a string. Each number represents one typewriter space ($.20^\circ$ of visual angle) and negative numbers refer to positions in the left visual field. A string which begins at -4 symmetrically spans the center of the fovea. As is illustrated, the number of letters reported peaks at about four letters for displays which span the center of the fovea and decreases to about two letters for displays somewhat in the periphery. If short-term memory capacity were limiting performance in full-report procedure, one would probably expect a flat function, since the capacity limit would provide a ceiling effect. In other words, if the subjects were actually "seeing" many more letters than were reported, one would ex-

pect the number of letters reported to equal short-term memory capacity over the small range of the retina considered here, rather than the smooth decline which actually occurs as the display is moved from the center of the fovea.

Even though the arguments presented above are hardly conclusive, I have chosen to attempt to model the data from report paradigms without reference to memory limitations. The same set of processes is used in dealing with all of the paradigms considered.

THE EMPIRICAL BASE

None of the models presented above makes any quantitative predictions about the rather consistent serial position curves found in most report paradigms and in many detection paradigms. A serial position curve is a description of performance as a function of the position of an item in a display. A number of factors typically vary with an item's display position: the retinal location of the item, its position in the report and processing orders, and the number and position of surrounding letters.

The role of retinal location in letter identification has been somewhat unclear. There is no question that threshold acuity decreases very rapidly as one leaves the center of the fovea, declining to less than half of maximum acuity within a degree of the center (Alpern, 1962). There is no reason, however, to assume that the letters presented in typical tachistoscopic experiments are near threshold size or thickness. Several investigators have suggested that retinal location is not a particularly important variable in letter identification (Bryden, 1966; Crovitz & Schiffman, 1965). Other experimenters have suggested that retinal location may not be very important by itself, but that it can become important in interaction with string position. Both Estes and Wolford (1971) and White (1970) reported that only items in the middle string positions are affected by retinal locus.

As I mentioned above, several factors vary with display position, and most investigators have exercised little or no control over these variables in examining

retinal location. For instance, most investigators do not instruct their subjects as to the order in which to report the letters. If subjects report the letters in a mixture of orders in keeping with their biases, and if report order affects performance, it would be quite difficult to partial out the retinal effect from the various order effects. Several experiments have been carried out to examine the effects of retinal location with the other variables partialled out. One such experiment (Hollingsworth & Wolford, Note 1, Experiment 1) examined the effect of retinal location in a simple situation, with report order eliminated as a variable by having the subject report only one letter. Single consonants were presented at one of five retinal locations flanked by zero, one, or two noise letters (the letter U). Subjects were instructed to report only the consonant on each trial. The results of primary concern for this article are presented in Figure 2. There are several important points to notice in the data. The first is that even in the absence of interference, letter identification performance did decrease with increasing distance from the center of the fovea. The decrease was not nearly as steep as the threshold function in Alpern (1962), but, as mentioned earlier, there is no reason to believe that the stimuli approached threshold size or thickness. The second important point is that the mere presence of unreported adjacent letters decreased performance *and* accentuated the retinal locus effect.

In another experiment retinal location and string position were varied orthogonally (Wolford & Hollingsworth, 1974c, Experiment 1). Nine-letter linear arrays were presented at 18 different retinal locations. Subjects were instructed to report as many letters as possible and to process and report the letters in a left-to-right order. Their adherence to the instructions was monitored and feedback was presented every 15 trials. With this design it is possible to look at the effect of retinal locus with string position held constant. For instance, the leftmost letters of each string occupied the same position in the report and processing orders and had the same

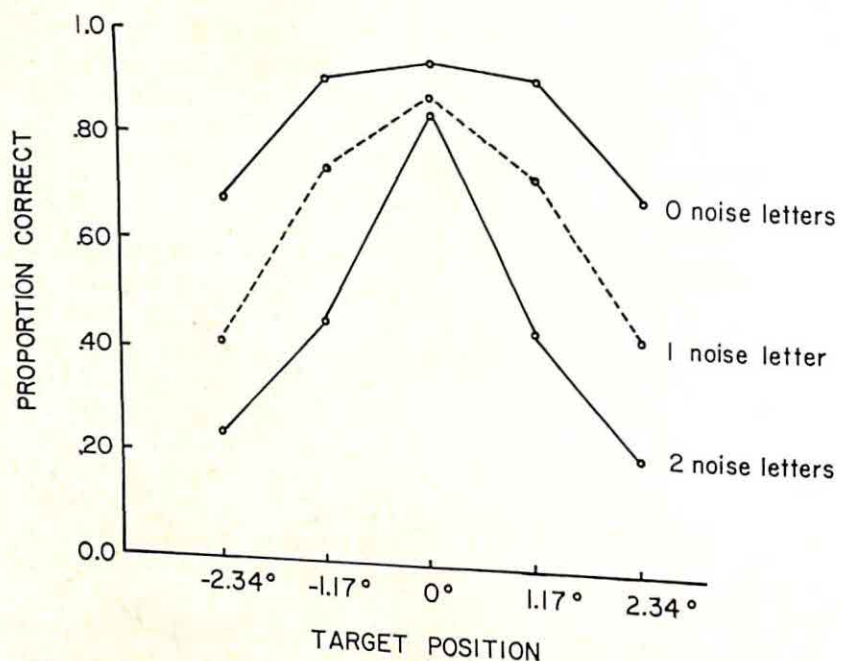


FIGURE 2. Letter recognition as a function of retinal position for three levels of lateral masking (from Hollingsworth & Wolford, Note 1).

number of surrounding letters. Any variation in performance among the leftmost letters, therefore, must be attributable to retinal location. In a similar manner the retinal function can be examined for each of the nine string positions. Several such functions are presented in Figure 3. Performance along any given curve is solely a function of retinal locus, and differences among curves are a function of string position. String position in this experiment was a composite of the effects of processing order, report order, and spacing. It is apparent in Figure 3 that retinal location had an effect at all string positions but that the effect was especially marked at the middle positions. Except for the first and last string positions, the retinal functions appeared quite regular. They all peaked at the center of the fovea and decreased rapidly over the first 1.5° from the center. It is also clear that when retinal location was held constant string position was a powerful variable. Performance decreased rapidly over the first several positions and rose somewhat at the end of the string. The two end positions (1 and 9) were somewhat less affected by retinal locus than the other positions. One of the distinctive features of the ends is the pres-

ence of a space on one side. Several investigators have demonstrated that letters receive interference from nearby adjacent letters (Estes & Wolford, 1971; Shaw, 1969; Townsend, Taylor, & Brown, 1971). A space on one side of a letter, therefore, potentially frees that letter from some amount of interference. Two properties of that interference were of particular importance in the development of the model to be presented in the next section. The first is that the amount of interference increases with increased distance into the periphery, and the second is that the interference appears to be asymmetric, with letters receiving more interference from letters on their peripheral side than from letters on their foveal side (Wolford & Hollingsworth, 1974b). In the Wolford and Hollingsworth (1974b) experiment, subjects were presented with nine-letter linear arrays of consonants at five different retinal locations. Subjects were instructed to process and report the letters in a left-to-right order, and spaces were introduced into some of the strings either before or after the fifth letter in the instructed processing order (pre- and postspaces). The performance on the fifth letters in that experi-

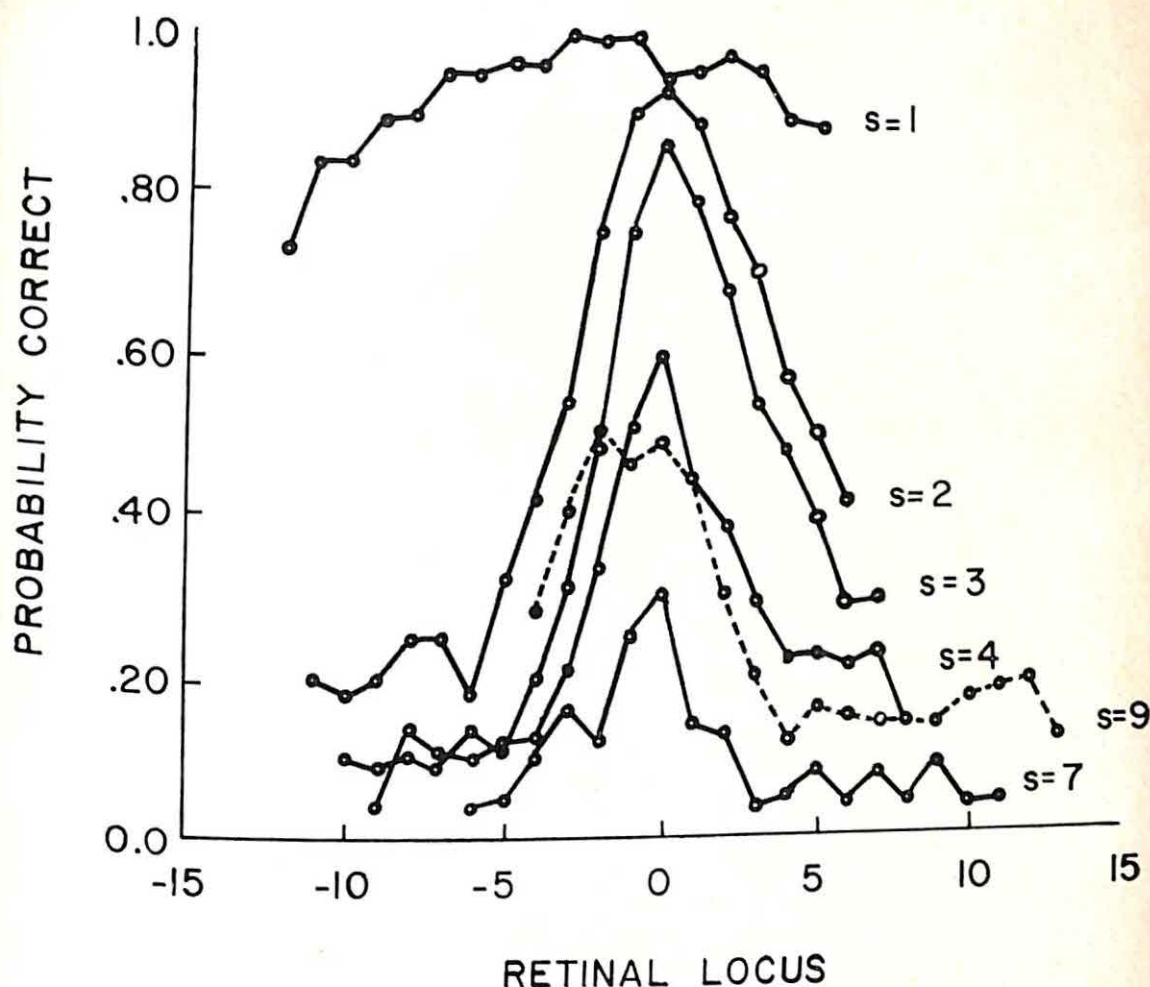


FIGURE 3. Performance as a function of retinal location at each of several string positions, denoted by s . (From "Retinal Location and String Position as Important Variables in Visual Information Processing" by G. Wolford and S. Hollingsworth, *Perception & Psychophysics*, 1974, 16, 437-442. Copyright 1975 by The Psychonomic Society, Inc. Reprinted by permission.)

ment is shown in Figure 4. Some of the displays contained two A s before or after the fifth letters instead of spaces. Subjects were instructed not to report those A s. The purpose of the A displays was control of the report order advantage provided the fifth letter by prespaces but not by postspaces. In other words, in addition to any unmasking effect that a space might have, a space before a letter in the instructed processing order also moves that letter up in the processing order. The sensory effect of a space can be examined by comparing each space condition with its corresponding A condition. The numbers on Figure 4 refer to the t values between the appropriate comparisons. When the fifth letter was at the center of the fovea, a space on either side was equally ineffective. When

the fifth letter was in the periphery, spaces had a marked effect. As mentioned earlier, however, the space on the peripheral side of a letter was much more effective than a space on the foveal side. In the left visual field, then, a prespace was more effective, while in the right visual field, a postspace was more effective. The increased performance at String Positions 1 and 9 in Figure 3 and the flatness of the retinal function at those positions are probably due, in part, to such spacing effects.

Even with the spacing effects held relatively constant, there was still a strong string-position effect in Figure 3. This effect is indicated by the consistent drop in performance as one moves among the middle string-position curves. These positions each had at least one letter on both

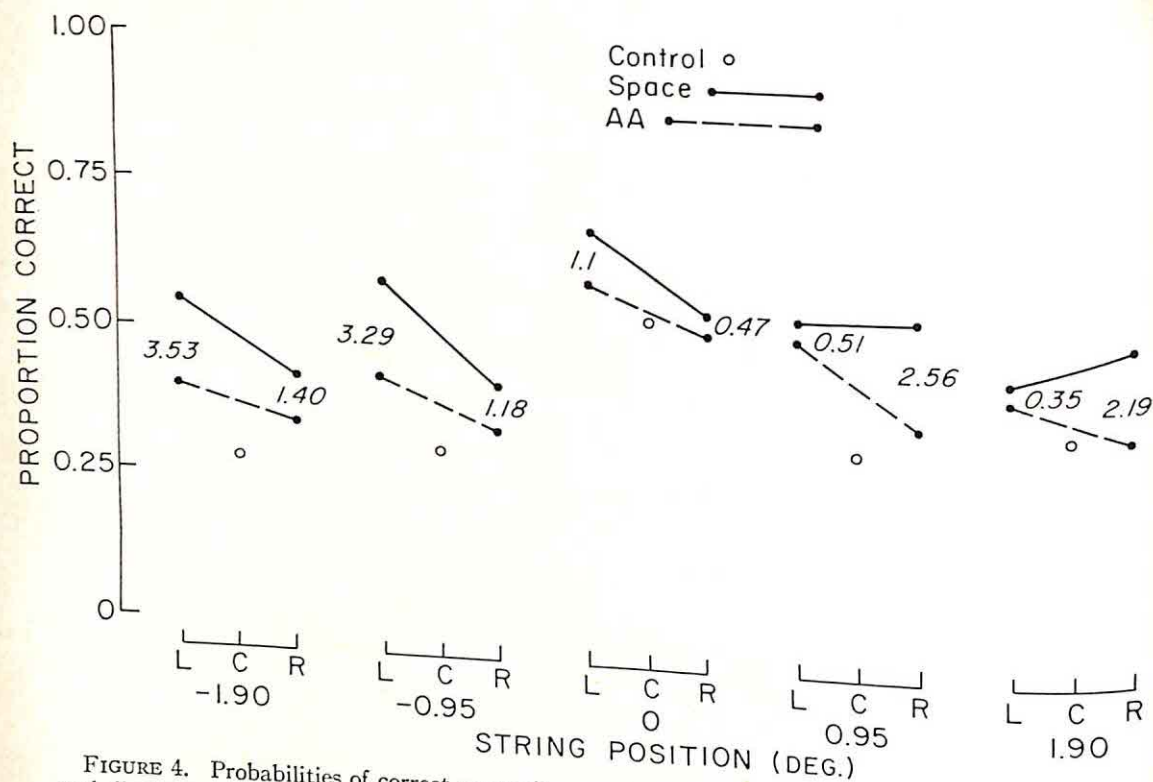


FIGURE 4. Probabilities of correct responding on the fifth letters as a function of retinal locus and display type. String positions are measured in degrees. On the abscissa, the L refers to a space, or AA, on the left of the fifth letter, R refers to a space, or AA, on the right of the fifth letter, and C refers to displays without spaces, or AA. (From "Lateral Masking in Visual Information Processing" by G. Wolford & S. Hollingsworth, *Perception & Psychophysics*, 1974, 16, 315-320. Copyright 1974 by The Psychonomic Society Inc. Reprinted by permission.)

sides. The performance drop across these positions was probably due to variation in the report and processing orders, that is, items farther out on the report and processing orders were identified less accurately.

There is some question as to whether a processing order exists that is conceptually independent of report order. In the experiment depicted in Figure 3 the two orders were intentionally confounded. To examine the effect of processing order by itself, two additional experiments were conducted (Wolford & Hollingsworth, 1974c, Experiments 2 and 3) in which processing order was varied but report order was held constant. The results of those experiments indicated that processing order was a significant variable in its own right (i.e., performance decreased with increasing processing order even when report order was held constant).

To return to the discussion at the beginning of the empirical section—the concern was with the identification of the variables

that determine the shape of the serial position curves. The shape of a particular curve depends upon the retinal placement of the array, the shape of the array (linear, matrix, etc.), and the instructions (or lack of instructions) given to the subjects concerning the report and processing orders. When no instructions are given, English-speaking subjects appear to have a strong bias to process in a left-to-right order (Bryden, 1966). They are capable, however, of processing in other directions (Estes & Wolford, 1971). If no order instructions were given and linear arrays were presented which spanned the center of the fovea, subjects might be expected to employ a mixture of orders. Serial position curves, therefore, should be easier to interpret when report and processing orders are controlled. Three such serial position curves from different parts of the retina are presented in Figure 5 (Wolford & Hollingsworth, 1974c, Experiment 1). One was chosen from the left visual field, one spanned the center,

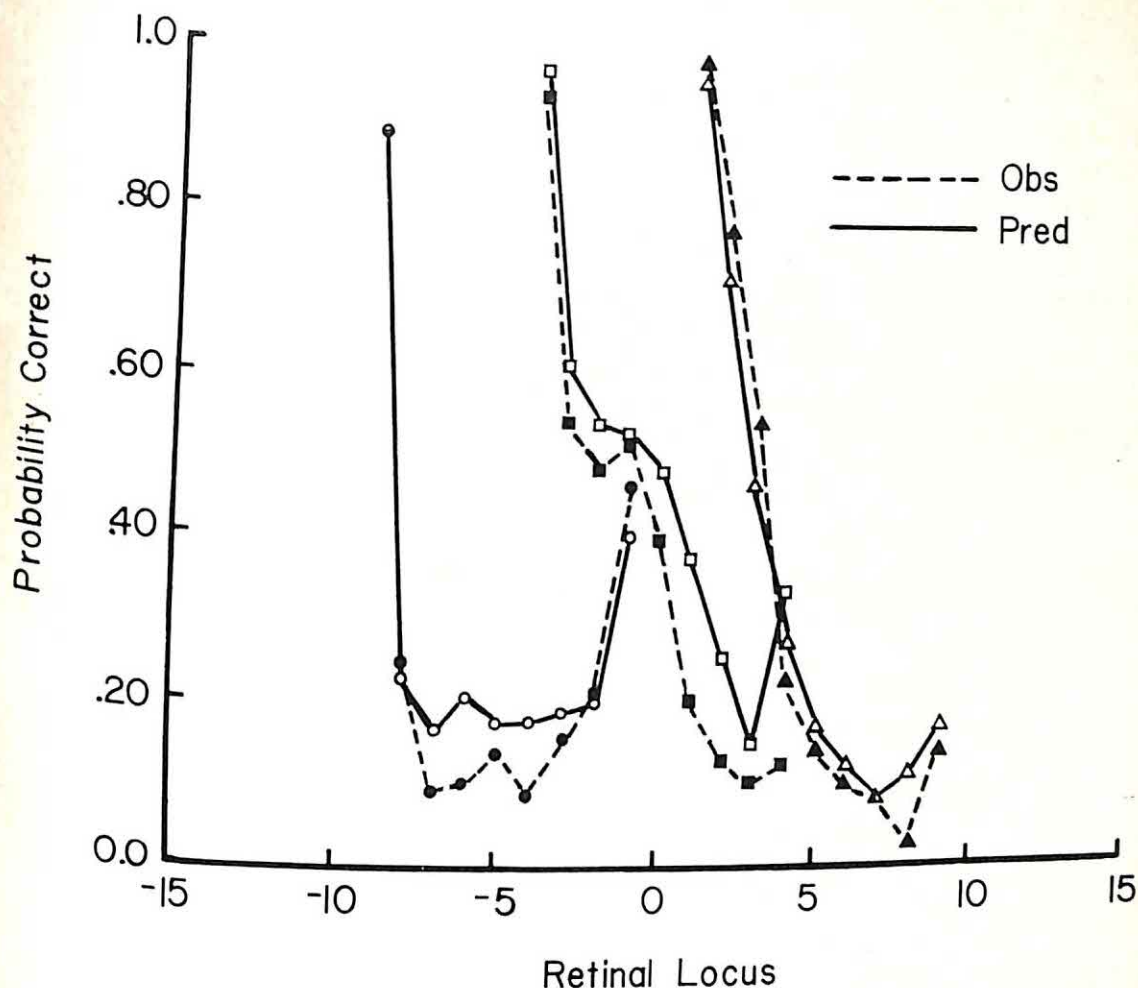


FIGURE 5. Observed and simulated serial position curves as a function of the retinal placement of a display. (Abbreviations: Obs = observed; Pred = predicted.)

and one from the right visual field. Subjects were instructed to process and report in a left-to-right order. These curves illustrate some combinations of the variables described above (e.g., the interaction of retinal location, processing order, and spacing for a particular string).

THE PERTURBATION MODEL

The perturbation model was formulated to provide a quantitative account of the data presented in the preceding section. In various assumptions the model borrows freely from the models described earlier. The section begins with a listing of the assumptions of the model, followed by a number of explanatory comments concerning the assumptions. Finally, a number of details involved in the actual simulation of the model will be discussed.

Extraction Assumptions

1. Information is extracted from visual displays at the feature level via independent parallel channels and is then stored in a sensory information store.

2. The probability of a feature being extracted is a function of retinal location, size, contrast, etc.

3. The spaces between letters are extracted just like other features, namely, as a function of their retinal location, width, etc.

Processing and Decision Assumptions

4. The information in the sensory store is processed in serial in order to identify the letters. Subjects have considerable control over their processing order.

5. The first task of the processor is to parse the various features into groups. A

feature group is defined as all of the non-space features between two spaces.

6. The decision criteria used in identifying the letters depend upon the task. For most of the report tasks considered here, it is assumed that if all of the features of a letter and no additional features are present in a feature group, then that letter is output. Additionally, it is assumed that if the features in a group represent a proper subset of one and only one letter, then that letter is output. Otherwise, an omission error is made.

7. The time it takes to process a letter is a function of the amount of information required. Different experimental paradigms require different amounts of information per letter. For simplicity, it is assumed that each letter in a report task requires one Δt of processing time.

Interference Assumptions

8. All of the extracted features, including spaces, are assumed to have some spatial ordering. Perturbations in that ordering occur with some probability during each Δt of time, that is, changes occur in the relative spatial ordering of the features over time.

9. The probability of a perturbation of a particular feature is a function of the retinal location of the feature and the amount of interletter spacing in the display. The wider the spacing in the display the smaller the probability of a perturbation between feature groups.

Some explanatory comments on these assumptions are clearly in order. The first is that Assumption 1 is similar to those proposed by Shiffrin and Geisler (1973). Second, letters interact with one another primarily due to perturbations during processing, and only to a lesser extent at the feature extraction stage, as in Estes' model. Some interaction between letters occurs at the extraction stage when there is a failure to extract a space. Third, although subjects are assumed to have control over their processing order, some orders are easier than others due to experience and/or the laws of physics (e.g., processing every other letter would probably be difficult). Fourth,

most theorists have not considered the problem of parsing features into groups and by default have assumed that features are automatically placed in appropriate groups. When spaces are not detected in the perturbation model, the features from two or more letters are placed in the same feature group and the identification process suffers, since it is unlikely (but not impossible) that the combined features from two or more letters would equal some other letter or a subset of some other letter.

The decision criteria described in Assumption 5 for report tasks are fairly conservative. For instance, if a particular feature group is a subset of two or more letters, an omission error will be made. One could assume that subjects would guess from the appropriate letters in that case rather than make an omission error. The conservative choice is based on the fact that most subjects in our report tasks make relatively few intrusion errors. Most of the more efficient decision rules produce many more intrusions. Different decision rules are specified in later sections for the confusion experiment and for the detection paradigm.

Some simplifications of the previous assumptions were used in the actual simulation of the model. The first was that all of the features in a letter were assumed to occupy the same retinal location in the display. Second, in implementing the perturbation process, spaces were considered fixed and all nonspace features were free to move at most one feature group per Δt . The perturbation process then becomes a random walk, where the states are represented by the various feature groups. The probability of a transition between two feature groups is a function of the retinal location of the two feature groups.

The actual implementation of the model also involved the choice of a feature space and the necessary mathematical functions. For the feature space, the table of features by letters from Lindsay and Norman (1972, p. 120) was used. Their feature table consists of seven types of features, such as vertical lines, acute angles, and continuous curves. A letter is characterized by a certain number of each type of feature.

The number of features per letter ranges from 1 for the letter *C* to 10 for the letter *B*. The "true" feature space in a particular experiment clearly depends upon the particular font used in that experiment. The feature set from Lindsay and Norman was chosen for this simulation because it seemed to provide a reasonable characterization of the font used in the experiments. One inadequacy of the feature set is the lack of any relational information. For instance, no information is included about the relative placement of the features within the letters.

Several functions were required for the simulation of the model. These functions, which determine feature extraction and perturbation, are not directly observable and are difficult to approximate from other considerations. The functions chosen were all exponential, and no attempt is made to defend the particular choice. Other sets of functions probably exist that would produce similar results, but the primary concern in this article lies in the specification of the processes involved and not in the choice of the functions. The actual simulation was carried out to demonstrate that there exists at least one set of functions and parameters that provide a reasonable approximation of the empirical data in the context of the above assumptions.

The following functions are required for the implementation of the model: (a) The probability of a feature being extracted from a character x° from the center of the fovea is

$$ae^{-(x^3/b)}, \quad (1)$$

where $x > 0$. The cubic, while somewhat inelegant, was required to account for the wide plateau in the single letter data around the fovea. (b) The transition probability for a feature moving from a feature group x° from the center to a group nearer the fovea is

$$k_1 e^{-(x^2/h_1)}. \quad (2)$$

(c) The transition probability for a feature moving from a group x° from the center to a group farther from the fovea is

$$k_2 e^{-(x^2/h_2)}. \quad (3)$$

Since the perturbation process is assumed to be a random walk, the sum of

Equations 2 and 3 must be less than or equal to 1.0 for a given value of x . The following parameter values were used in all of the subsequent simulations: $a = .995$, $b = 160$, $k_1 = .988$, $h_1 = 64$, $k_2 = .988$, and $h_2 = 16$. The parameters a and b would be expected to vary with the physical properties of the display, and in particular with the size and thickness of the characters used. The h s and k s would depend primarily upon the amount of spacing between the letters in a display. In the various simulations a , k_1 , and k_2 were quite sensitive. Small changes in their values had large effects on the simulated output. The parameter values used do not necessarily represent the optimal values. The simulation required about 45 minutes of computer time per 1,000 trials, so no attempt was made to exhaustively search the entire parameter space.

A sample run through a trial of the simulation might prove useful in conveying the nature of the model. This sample trial is representative of a simulation of the data depicted in Figure 3. The trial begins by choosing 1 of 18 retinal locations in which to present the display. Nine letters are then chosen at random from the consonants, without replacement, and the order of the letters is randomized. With some probability (specified by Equation 1 above), each of the features from each of the presented letters (including the spaces between the letters) is stored in a sensory register. For instance, if *B* were one of the presented letters, each of its 10 features would have an independent probability of being extracted, depending upon the location of *B* on the retina. The parser then places all of the nonspace features into feature groups, with a group being all of the nonspace features between two spaces. In accordance with the experimental instructions, the processor attempts to identify the letter represented by the leftmost feature group, using the decision rule stated in Assumption 6. Following the processing of that feature group, each of the features in all of the groups has some probability (as specified by Equations 2 and 3) of moving to an adjacent feature group. The processor then moves to the

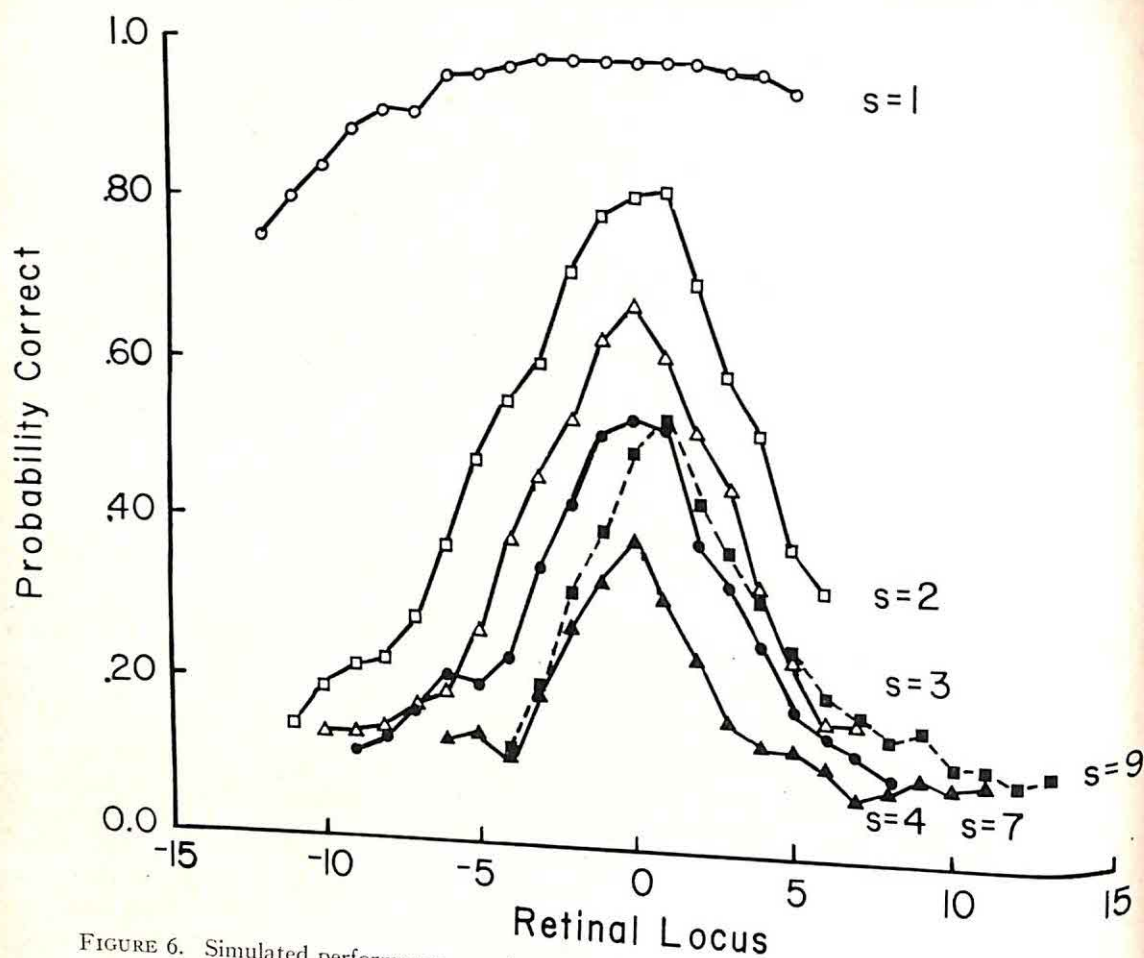


FIGURE 6. Simulated performance as a function of retinal locus at each of several string positions, denoted by s .

next group in the left-to-right order and attempts to identify it. The simulation's response on that trial would consist of those consonants for which a perfect match was found in one of the feature groups or for which a proper subset was found in one of the groups. Errors would consist primarily of omissions, although some intrusions would be made. Even a perfect match does not guarantee a correct response, since perturbations or unextracted features might have transformed the feature group, resulting from a presented letter into the perfect match of another letter.

ACCORDANCE WITH THE EMPIRICAL BASE

The model was first applied to the data from Wolford and Hollingsworth (1974c, Experiment 1; see Figures 3 and 5). That experiment had 162 data points (18 retinal locations crossed with 9 string positions).

The simulated values are based on 1,000 trials per point. One indication of the fit of the model is illustrated in Figure 5. The open points and solid lines represent the simulated values. The fit to that experiment is further illustrated in Figure 6, which corresponds to the empirical data presented in Figure 3. While it is difficult to illustrate the correspondence between two sets of 162 points simply, it appears that the model captures most of the functional nature of the data from that experiment quite well. In addition, there do not seem to be any strong systematic differences between the empirical and simulated data.

The final analysis from Wolford and Hollingsworth's (1974c) experiment that is relevant to the model concerns the mean number of letters output per nine-letter string as a function of the retinal location of that string. The significance of those

TABLE 1

SIMULATED PROPORTION CORRECT AS A FUNCTION OF RETINAL LOCUS AND SPACING

Retinal locus																		
-9	-8	-7	-6	-5	-4	-3	-2	-1	0	1	2	3	4	5	6	7	8	9
89	23	17	21	18	18	19	20	40										
86	29			35	19	16	22	38										
87	24	19	17	23			43	42										
					97	61	54	52	48	37	25	15	33					
					97	66			62	35	26	21	32					
					96	61	52	51	63			20	33					
										98	71	46	27	17	12	08	11	17
										97	84			17	09	08	10	13
										98	76	45	23	47			19	18

data were discussed earlier, and both the empirical and simulated data are shown in Figure 1. Once again there is reasonable correspondence between the observed and simulated values.

It might be useful to provide an intuitive explanation of why the perturbation model provides a reasonable fit to the empirical data. The letters at String Position 1 are not subject to much interference, since they are at the beginning of the processing order. Performance on the leftmost letters, therefore, reflects primarily the retinal extraction function presented as Equation 1. Performance decreases with string position because letters farther out on the processing order have undergone more perturbation by the time the processor reaches them. When string position is held constant, performance decreases with distance into the periphery, because the probability of a perturbation is higher in the periphery, and because those letters are less likely to have had all of their features extracted in the first place. Performance is elevated at String Position 9, because that letter is subject to perturbations on only one side.

The next step in evaluating the viability of the model was to examine the effects of introducing spaces into the strings. Spaces were introduced by removing the two letters preceding the fifth letter or the two letters subsequent to it. Processing was assumed to take place in a left-to-right direction. As was discussed in the empirical section, prespaces provide a report order advantage to the fifth letter that post-

spaces do not. As an empirical control for this discrepancy, unreported *As* were introduced into some strings. It is not clear how best to simulate the effect of an unreported *A* in the model. As a solution, the fifth letter on a prespace trial was assigned the same position in the processing order as the fifth letter on a postspace trial. The simulated space effects are presented in Table 1. Except for the prespace data, these curves can be compared with roughly similar data in Estes and Wolford (1971) or in Wolford and Hollingsworth (1974b). The prespace curves should differ slightly because of the absence of the report order advantage in the simulation.

The effects of spaces in the simulation were similar to those in the empirical data. Spaces near the fovea were relatively ineffective, although they were more effective than in the empirical data. Spaces were more effective away from the center, and spaces on the peripheral side of a letter were superior to spaces on the foveal side. The intuitive reasons for the spaces effects in the simulation are straightforward. Perturbations, by design, are more frequent in the periphery and are more likely to occur in the direction of the fovea since h_1 was considerably larger than h_2 . In addition, with the decision rule used, perturbations to a feature group are more damaging than perturbations from a group. Losing features preserves the group as a subset of the letter, while adding features to a group often destroys the subset relationship.

TABLE 2

VALUES FOR PROPORTION CORRECT AS A FUNCTION
OF SERIAL POSITION IN THE CONFUSION
EXPERIMENT

Values	Serial position				
	1	2	3	4	5
Observed	94	90	76	54	67
Simulated					
New rule	99	91	70	50	57
Old rule	97	76	45	24	36

The simulation was also applied to the report paradigm described in Wolford and Hollingsworth (1974a). Subjects were presented with five-letter consonant arrays beginning at the center of the fovea and extending into the right visual field. Subjects were instructed to process in a left-to-right order and to always output five consonants on every trial. The simulated and observed serial position curves for that experiment are presented in Table 2. The letters did not have to be output in the correct order to be scored as correct. The same parameter values were used as in the previous simulations. The decision rule, however, was altered, since the subjects were forced to output a letter for every letter presented. The new rule was as follows: If the features in a group matched a letter perfectly, that letter was output; if the features in a group were a subset of one or more of the consonants, the letter with the highest percentage of features in common with the group was output; otherwise a consonant was chosen from among all consonants with the same percentage of features in common. Table 2 also contains the simulated values, with the conservative decision rule used in the previous simulations. As is evident in Table 2, the new decision rule provides a much better fit to the data than the conservative rule.

In order to examine the types of errors made in the Wolford and Hollingsworth (1974a) experiment, confusion matrices were constructed from the errors for both the observed and simulated data. Those matrices were then clustered, using Johnson's (1967) nonmetric hierarchical cluster-

ing technique. All of the letters that clustered together above the fourth smallest diameter for the two clustering outputs are presented in Table 3. Letters that appear on the same row in Table 3 tend to appear as errors for each other in the data. Although there are many differences between the two outputs, there is a general similarity and a predominance of visual confusions in both cases. That similarity reflects primarily on the choice of the feature space and the decision rule.

POTENTIAL EXTENSION OF THE MODEL TO OTHER PARADIGMS

The perturbation model was specifically formulated to provide a quantitative account of the data from tachistoscopic experiments using the full-report procedure. The correspondence between the predicted and observed data, while by no means perfect, is sufficient to warrant further attention. There exist, however, numerous paradigms other than the report paradigms considered here. In fact, much of the recent literature has tended to focus on detection paradigms. Regardless of how well the perturbation model succeeds in fitting the data from report paradigms, it would be less than fruitful to invest further labor in it if the model were grossly inconsistent with the data obtained from other paradigms. In this final section, therefore, an attempt is made to show that the model is potentially consistent with data from other paradigms.

Sperling (1959) developed a partial report paradigm in which subjects were presented with displays containing three rows of letters. A tone was sounded subsequent to the offset of each display, indicating

TABLE 3
SUMMARY OF CLUSTERING RESULTS FOR ERRORS IN
THE CONFUSION EXPERIMENT

Simulated errors	Observed errors
BDFPR	FPRTXZ
HLT	HLJN
KQVWX	BDQSV
CGJS	CG
MNZ	MW

which of the rows was to be reported. Performance was facilitated compared to a toneless control condition, but only if the tone followed the display by less than a second. The perturbation model predicts similar results. Subjects are assumed to have control over their processing order. When the tone is sounded in the partial report task, then, the subject could immediately begin processing the indicated row if he was not already doing so. If the tone was sounded soon enough, the indicated row would be moved up in the processing order on the average and performance would be facilitated. If the delay was too great, processing would have been completed and the tone would have no effect.

In a detection task the subject is asked to identify which of k previously specified target letters appeared in a display containing n elements. Performance in detection tasks is generally superior to performance in full-report procedure, when measured as in Estes and Taylor (1964). The perturbation model also predicts superior detection performance for two reasons: A more efficient decision rule can be used, and the processor can move more rapidly. The preferred decision rule is similar to the rule used by Shiffrin and Geisler (1973). If a perfect match is encountered between a feature group and one of the targets, then that target is output. If no perfect match occurs, but a subset of one and only one of the targets is found among the feature groups, then that target is output. In all other cases, subjects choose randomly from the appropriate targets. It seems reasonable to assume that the processor moves more rapidly in a detection task, since it only has to compare each display element with the target letters rather than with the entire alphabet as in a report task. Estes (1972) found that latencies increased with increases in the size of the target set. The more rapidly the processor moves through a display the less the amount of perturbation; hence improved performance. For the same reason one would expect flatter serial position curves in detection tasks than in report tasks.

As mentioned earlier, Estes (1972) found that detection performance did not vary with display size if nonconfusable noise letters were used. Two letters would presumably be confusable only if they shared common features. Performance did decrease with display size with confusable noise items. Performance in the perturbation model decreases with increases in the number of confusable noise elements, because the target is moved farther out on the processing order on the average, and there is increased opportunity for one of the noise items to be confused with the incorrect target. Performance, however, does not decrease as nonconfusable items are added. The target is still moved farther out on the processing order on the average, but since features *move* over time rather than decay, performance does not deteriorate. If a subject finds one of the target features in a feature group, he knows that it came from one of the targets, since the noise letters by definition had no such features. Retinal location would be the only variable affecting performance in the nonconfusable case. If enough noise elements were used so that the target was often moved far into the periphery, performance might deteriorate. The retinal extraction function, Equation 1, is quite flat, however, in the area covered by the displays in Estes (1972).

Gardner (1973) designed an experiment to distinguish between his unlimited-capacity model and Rumelhart's (1970) limited-capacity model. Subjects were presented with two detection tasks on each trial, represented by the two rows of a display. Difficulty within a row was varied by the presence or absence of a confusable noise element. Neither the noise nor target elements between rows were confusable. All elements were equidistant from the center of the fovea and were separated by more than a degree of visual angle. Gardner's choice of spacing was based on the finding that letters separated by a degree or more do not appear to interact (Collins & Eriksen, 1967). Gardner's main finding that performance in one of the tasks was not affected by the level of difficulty in the other task was

claimed to be inconsistent with plausible attention distributions in Rumelhart's model. The perturbation model could yield predictions similar to Gardner's results. Perturbations between feature groups are a function of distance and could be assumed to approach zero with his choice of spacing. Processing order would not matter then, and the difficulty in one task would have no bearing on the other. Difficulty within a task affects performance, due to an increased probability of a confusion with the incorrect target. In essence, the perturbation model reduces to Gardner's independent parallel-channels model when retinal locus is held constant and the spacing between letters is sufficiently wide.

As mentioned before, processing order is not a relevant variable when spacing is sufficiently wide to prevent perturbations. For this reason the model is also consistent with the finding by Shiffrin and Gardner (1972) that sequential exposure was not superior to simultaneous exposure in their detection task. When spacing is wide, the perturbation model behaves very much like an unlimited-capacity model. When spacing is close, processing order serves a role similar to the distribution of attention in Rumelhart's model. Decay is time-based in Rumelhart's model but depends upon both time and distance in the perturbation model.

There are findings which are inconsistent with the perturbation model as currently formulated. One such finding is that the percentage of transposition errors is lower in the model than in the real data. In Wolford and Hollingsworth (1974a), approximately one-third of the errors were transpositions. The perturbation model, in simulating that experiment, made fewer transposition errors. Another troublesome finding comes from an experiment by Townsend et al. (1971). They presented subjects with linear arrays at long exposure durations (up to 3 seconds) while preventing eye movements. The serial position curves in their data were quite similar to the ones presented earlier. One would probably expect the feature extraction process in the perturbation model to continue

throughout the exposure duration and yield a fairly flat serial position curve. Any resolution of the model with these data will have to await further research into the effects of long exposure durations sans eye movements. There is also no mechanism, at present, for handling the effects of context, which can be quite powerful in visual information processing (Mewhort, 1967; Rumelhart & Siple, 1974).

The last issue to be considered is the relationship of the perturbation model to other existing models and to other potential formulations. As mentioned earlier, none of the quantitative models discussed previously has been formulated to deal with serial position effects in report experiments. If one deals with detection paradigms, the perturbation model would have much in common with all three models, but would be particularly similar to the models developed by Gardner (1973) and Shiffrin and Geisler (1973).

The primary difference between the perturbation model and earlier quantitative models is the addition of a serial processor and the specification of the perturbation process. A serial processor has been postulated by a number of other investigators and was considered necessary to account for the potent effects of report and processing order. It may be possible to formulate a parallel processor which works as well or better, but, at present, it is not clear to us how to do so. The perturbation process was chosen instead of a decay process so that the model would be potentially consistent with many recent detection experiments (Collins & Eriksen, 1967; Gardner, 1973; Shiffrin & Gardner, 1972). As was discussed previously, we chose to model the various report paradigms without reference to capacity limits in short-term memory. Some interaction with short-term memory may be required to account for the number of transposition errors in our report experiments, but any very potent role would be difficult to reconcile with the data presented in Figure 1.

The fact that the perturbation model is qualitatively consistent with many of the paradigms considered here in no way im-

plies that it could provide adequate quantitative accounts of the data from those paradigms. Most of those paradigms used physical parameters and display shapes that were quite different from the empirical base described in this article, and different parameters and perhaps functions would be required to attempt such fits. The model does provide a fairly good account of the report paradigms considered here and seems generally consistent at a qualitative level with many other paradigms. Further research is clearly in order to obtain more direct evidence for some of the processes postulated in the model.

REFERENCE NOTE

1. Hollingsworth, S. R., & Wolford, G. *Lateral masking and retinal position effects in tachistoscopic processing*. Unpublished manuscript, 1974. (Available from G. Wolford, Dartmouth College, Hanover, N. H. 03755.)

REFERENCES

- Alpern, M. Muscular mechanisms. In H. Davson (Ed.), *The eye* (Vol. 3). New York: Academic Press, 1962.
- Bryden, M. P. Accuracy and order of report in tachistoscopic recognition. *Canadian Journal of Psychology*, 1966, 20, 262-272.
- Collins, J. F., & Eriksen, C. W. The perception of multiple simultaneously presented forms as a function of foveal spacing. *Perception & Psychophysics*, 1967, 2, 369-373.
- Conrad, R. Acoustic confusion in immediate memory. *British Journal of Psychology*, 1964, 55, 75-84.
- Crovitz, H. F., & Schiffman, H. R. Visual field and the letter span. *Journal of Experimental Psychology*, 1965, 70, 218-223.
- Estes, W. K. Interactions of signal and background variables in visual processing. *Perception & Psychophysics*, 1972, 12, 278-286.
- Estes, W. K. Phonemic encoding and rehearsal in short-term memory for letter strings. *Journal of Verbal Learning and Verbal Behavior*, 1973, 12, 360-372.
- Estes, W. K. Redundancy of noise elements and signals in visual detection of letters. *Perception & Psychophysics*, 1974, 16, 53-60.
- Estes, W. K., & Taylor, H. A. A detection method and probabilistic models for assessing information processing from brief visual displays. *Proceedings of the National Academy of Sciences of the United States of America*, 1964, 52, 446-454.
- Estes, W. K., & Wolford, G. L. Effects of spaces on report from tachistoscopically presented letter strings. *Psychonomic Science*, 1971, 25, 77-79.
- Gardner, G. T. Evidence for independent parallel channels in tachistoscopic perception. *Cognitive Psychology*, 1973, 4, 130-155.
- Johnson, S. C. Hierarchical clustering schemes. *Psychometrika*, 1967, 32, 241-254.
- Kolers, P. A., & Rosner, B. S. On visual masking (metacast): Dichoptic observation. *American Journal of Psychology*, 1960, 73, 2-21.
- Lindsay, P. H., & Norman, D. A. *Human information processing*. New York: Academic Press, 1972.
- Matthews, M. L. Locus of presentation and the selective masking effect. *Canadian Journal of Psychology*, 1973, 27, 343-349.
- Mewhort, D. J. K. Familiarity of letter sequences, response uncertainty and the tachistoscopic recognition experiment. *Canadian Journal of Psychology*, 1967, 21, 309-321.
- Rumelhart, D. E. A multicomponent theory of the perception of briefly exposed visual displays. *Journal of Mathematical Psychology*, 1970, 7, 191-218.
- Rumelhart, D. E., & Siple, P. Process of recognizing tachistoscopically presented words. *Psychological Review*, 1974, 81, 99-118.
- Shaw, P. Processing of tachistoscopic displays with controlled order of characters and spaces. *Perception & Psychophysics*, 1969, 6, 257-266.
- Shiffrin, R. M., & Gardner, G. T. Visual processing capacity and attentional control. *Journal of Experimental Psychology*, 1972, 93, 72-82.
- Shiffrin, R. M., & Geisler, W. S. Visual recognition in a theory of information processing. In R. L. Solso (Ed.), *Contemporary issues in cognitive psychology: The Loyola symposium*. Washington, D.C.: V. H. Winston, 1973.
- Sperling, G. *Information in a brief visual presentation*. Unpublished doctoral dissertation, Harvard University, 1959.
- Sperling, G., & Speelman, R. G. Acoustic similarity and auditory short-term memory: Experiments and a model. In D. A. Norman (Ed.), *Models of human memory*. New York: Academic Press, 1970.
- Townsend, J. T., Taylor, S. G., & Brown, D. R. Lateral masking for letters with unlimited viewing time. *Perception & Psychophysics*, 1971, 10, 375-378.
- White, M. J. Retinal locus and the letter-span error function. *Perception & Psychophysics*, 1970, 8, 107-109.
- Wolford, G., & Hollingsworth, S. Evidence that short-term memory is not the limiting factor in the tachistoscopic full report procedure. *Memory & Cognition*, 1974, 2, 796-800. (a)
- Wolford, G., & Hollingsworth, S. Lateral masking in visual information processing. *Perception & Psychophysics*, 1974, 16, 315-320. (b)
- Wolford, G., & Hollingsworth, S. Retinal location and string position as important variables in visual information processing. *Perception & Psychophysics*, 1974, 16, 437-442. (c)

(Received June 10, 1974;

revision received January 3, 1975)

Role of the Hypothalamus in the Regulation of Food and Water Intake

Sebastian P. Grossman
University of Chicago

A review of recent experimental findings suggests that the classic model of hypothalamic control of simple appetitive motivational states may be based, at least in part, on possibly faulty interpretations of the effects of lesions in the ventromedial or lateral hypothalamus on food and water intake. It appears increasingly probable that such lesions may affect ingestive behavior because they interrupt fibers of passage, which may or may not synapse on local neurons. In the case of the lateral hypothalamic lesion syndrome, fiber systems that are part of the complex feedback system that constitutes the extrapyramidal motor system have been implicated. This raises the possibility that sensorimotor dysfunctions may play a more important role in the aphagia and adipsia than is commonly assumed, and some recent behavioral observations of rats with such lesions emphasize this possibility. In the case of the ventromedial hypothalamic lesion syndrome, there is evidence that the overeating and obesity may be related, at least in part, to a disruption of ascending noradrenergic pathways which may or may not synapse in the region, and that steroid hormone dysfunctions may be responsible for some of the abnormalities in ingestive behavior.

THE CONTEMPORARY MODEL OF HYPOTHALAMIC REGULATION

Little of significance was known about the brain mechanisms that regulate ingestive behavior in accordance with fluctuations in the organism's energy and fluid needs until Hetherington and Ranson (1940) demonstrated unequivocally that small bilateral lesions in the ventromedial hypothalamus produce overeating and obesity. Anand and Brobeck (1951) reported about 10 years later that lesions lateral and dorsal to the ventromedial area produced opposite, inhibitory effects on food and water intake.

More detailed analyses of the effects of these hypothalamic lesions have shown that they appear to be essentially irreversible. Rats with ventromedial hypothalamic lesions eventually reach a "static" phase and no longer eat more than their grossly enlarged bodies require. However, these animals return to a "dynamic" state of intense overeating when starved to normal body weights (Brobeck, 1946). Rats with lateral hypothalamic lesions often recover seemingly normal food and water intake when properly nursed through weeks and

often months of aphagia and adipsia, but their ingestive behavior does not appear to be under control of the regulatory influences which determine intake in the normal animal. These rats are adipsic when food deprived and do not respond normally to hyperosmotic or hypovolemic challenges that evoke drinking in the intact rat (Epstein & Teitelbaum, 1964; Stricker & Wolf, 1967). It is generally accepted that the water intake which occurs in these animals is related mainly to the ingestion of dry laboratory food (Kissileff & Epstein, 1969), although delayed reactions to experimentally induced hypovolemia and cellular dehydration have been observed (see Stricker, 1973). The feeding behavior of rats that seem to have recovered from bilateral lesions in the lateral hypothalamus is similarly unresponsive to stimuli which elicit food intake in the normal animal. They do not eat in response to insulin (Epstein & Teitelbaum, 1967) nor to the glucose analogue 2-deoxy-D-glucose (Wayner, Cott, Millner, & Tartaglione, 1971), and they overreact to dietary admixtures that change the taste or texture of their food (Teitelbaum & Epstein, 1962).

These observations suggest that the food intake of the recovered animal may be motivated primarily by palatability rather than nutritional needs.

These and related findings have given rise to the hypothesis that food intake may be regulated by a pair of interacting hypothalamic mechanisms. The most detailed version of such a model (Stellar, 1954) proposes that the motivation to eat may be proportional to the activity of a "feeding center" in the lateral hypothalamus and that the "satiety center" of the ventromedial hypothalamus influences feeding by exerting inhibitory effects on the lateral hypothalamus. Both mechanisms are thought to collect neural as well as chemical information about the nutritional state of the organism and to receive inputs not only from peripheral sources but also from other regions of the brain. Although the inputs and outputs of these hypothalamic centers were specified only vaguely, this hypothesis has enjoyed wide acceptance and has had considerable heuristic success in directing the research of many laboratories.

The results of this research have, in general, confirmed the hypothesis that the hypothalamus plays an important role in the regulation of food and water intake. However, the unique position which contemporary theories have accorded to it and, indeed, the nature of the influence it may exert have been questioned. The most fundamentally important issue raised by the results of recent investigations concerns the possibility that some or all of the effects of lesions in the hypothalamus may be due to an interruption of fibers which may not synapse in the area. A related area of concern is the possibility that the hypothalamic mechanisms may not perform unique integrative functions but represent merely nodal points in a broadly distributed system of mutually interacting, feeding- and drinking-related pathways. Since many of the diffuse fiber systems that course through the hypothalamus are important components of the extrapyramidal motor system and related sensory pathways, the possibility must be considered

more seriously than it has been in the past that sensory or motor influences may play a major role in the disturbances in feeding or drinking that are observed in animals with hypothalamic lesions.

A POSSIBLE ROLE OF EXTRAHYPOTHALAMIC MECHANISMS

The uniqueness of the hypothalamic mechanisms has been questioned in a flurry of papers reporting that lesions in several extrahypothalamic structures produce effects on food and water intake that are, in many cases, as severe, persistent, and apparently specific as those seen after hypothalamic damage. I reviewed this literature in 1972 (Grossman, 1972) and shall not attempt to duplicate this effort here. A few examples and a review of some recent developments may, however, be useful.

Complete and persistent aphagia and adipsia have been reported following lesions in the globus pallidus (Morgane, 1961a), midbrain tegmentum (Blatt & Lyon, 1968; Gold, 1967), and substantia nigra (Ungerstedt, 1971). Less persistent aphagia and adipsia have been seen after temporal lobe damage (Pubols, 1966) and after lesions in the lower brainstem or basal thalamus, which interrupt projections of the trigeminal nerve (Zeigler & Karten, 1974). Hyperphagia has been reported after lesions in the mammillary region (Graff & Stellar, 1962), frontal lobe (Richter & Hawkes, 1939), temporal lobe (Grossman & Grossman, 1963), thalamus (Schreiner, Rioch, Pechtel, & Masserman, 1953), and periaqueductal gray of the brainstem (Skultety & Gary, 1962), following chemical or surgical lesions in portions of the mesencephalon that are traversed by the ventral noradrenergic bundle (Ahlskog & Hoebel, 1973) and after surgical transection of the anterior (Grossman & Grossman, 1970), lateral (Albert & Storlien, 1969; Scalfani & Grossman, 1969), and posterior (Grossman & Grossman, 1970) connections of the medial hypothalamus (discussed later). Hyperdipsia has been seen after some brainstem lesions in the vicinity of the interpeduncular nucleus (Kent & Grossman, unpublished observation, 1974), after

surgical transection of the posterior fiber connections of the ventromedial hypothalamus (Grossman & Grossman, 1970), and after septal lesions (Harvey & Hunt, 1965).

There is some evidence that at least some of these structures may be arranged in series (to borrow an analogy from electronic circuits). For example, the mild and temporary effects of a unilateral lesion in the lateral hypothalamus are converted into severe and persisting aphagia and adipsia not only by a contralateral hypothalamic lesion but also by a unilateral lesion in the contralateral tegmentum (Gold, 1967) or globus pallidus (Krantz & Teitelbaum, cited in Teitelbaum, 1967). A similar interaction has been reported between the effects of lesions in the mid-brain and in the subthalamus and zona incerta (Blatt & Lyon, 1968). Additive effects of hyperphagia-producing asymmetric knife cuts and lesions have also been reported (e.g., Gold, Quackenbush, & Kapatos, 1972).

An observation by Walsh and Grossman (1973) suggests that the effects of lateral hypothalamic lesions on water intake, as well as the impaired or absent feeding response to cellular glucoprivation, may in fact be the result of damage to tissues or fibers of passage of the zona incerta rather than to the hypothalamus itself. We found that small lesions in the anterior aspects of the zona incerta, which do not involve tissues of the lateral hypothalamus, completely abolish regulatory drinking without affecting food intake. The effects of these lesions may have escaped observation to date because animals with zona incerta lesions drink nearly normal quantities of water when given ad-libitum access to dry laboratory chow and display adipsia only when food deprived. This duplicates the behavior of rats that have recovered from the effects of lateral hypothalamic lesions, but we have noted one interesting exception (Walsh & Grossman, 1974). Whereas rats with lateral hypothalamic lesions do not drink in response to either intracellular dehydration or extracellular hypovolemia (although they may make long-term ad-

justments in intake, Stricker, 1973; Stricker & Wolf, 1967), rats with zona incerta lesions show a selective impairment in their response to cellular dehydration and react normally to extracellular hypovolemia. A similarly pure regulatory deficit has been reported by Blass and Epstein (1971) in the rat and by Peck and Novin (1971) in the rabbit after bilateral lesions in the lateral preoptic region. In the rat some of these lesions selectively abolished the response to cellular dehydration; others inhibited drinking in response to extracellular hypovolemia as well.

Walsh and I have investigated the relationship between the zona incerta and adjacent tissues of the preoptic region, hypothalamus, and thalamus, which have all been implicated in food- or water-related regulatory functions. We found that knife cuts which sever the direct connections between the zona incerta and preoptic region reduced water intake during periods of food deprivation (although not as drastically as zona incerta lesions) but had no effect on drinking in response to systemic injections of hypertonic saline or polyethylene glycol. Cuts beneath the zona incerta, which disrupt its connections with the lateral hypothalamus, not only interfered with the response to cellular dehydration but also inhibited the drinking response to extracellular hypovolemia, which we found to be unaffected by zona incerta lesions. This suggests that the zona incerta as well as the lateral preoptic region may contain components which are important for appropriate reactions to extracellular and cellular dehydration. The fact that our lesions in the anterior portions of the zona incerta (as well as some of Blass and Epstein's lateral preoptic region lesions) selectively affected the response to cellular dehydration indicates further that in both areas the two mechanisms may be, to some extent, anatomically distinct. This is further supported by our observation that circular knife cuts above the zona incerta produced small but consistent inhibitory effects on drinking in response to extracellular hypovolemia but had no effect on water intake after hypertonic saline.

These experiments have also provided some evidence that the osmo- or sodium-receptors, which are believed to monitor the state of cellular hydration, may be located in the zona incerta and lateral preoptic region. Blass and Epstein (1971) reported that microinjections of hypertonic saline or hypertonic sucrose into some sites in the lateral preoptic area elicited drinking in sated rats, and that similar injections of distilled water reduced water consumption in deprived animals. Peck and Novin (1971) noted similar effects in the rabbit. Walsh and I observed that microinjections of hypertonic saline into the zona incerta elicited drinking in sated rats but found hypertonic sucrose to be an ineffective stimulus (Walsh & Grossman, 1974). This observation is congruent with Andersson's suggestion that central receptors, selectively activated by sodium, may contribute to the regulation of water intake (see Andersson, 1973, for a review of related materials).

Epstein, Fitzsimons, and Rolls (1970) reported that microinjections of the powerful renal dipsogen angiotensin into the anterior hypothalamus, preoptic area, and septal region produced drinking in sated rats, suggesting that this region of the brain might also mediate the central effects of this compound. However, Johnson (in press) has recently shown that small doses of angiotensin elicited drinking only at cannula placements that involved injury to the ventricular system (and thus permitted the escape of the compound into the ventricles). Viewed in conjunction with Simpson and Routtenberg's (1973) observation that destruction of the subfornical organ, which protrudes into the ventricular system, abolishes drinking responses to intracranial angiotensin, these findings suggest that the subfornical organ rather than the preoptic-anterior hypothalamic region may be the central site of angiotensin action. This interpretation is further supported by Simpson and Epstein's (Note 1) recent report that subfornical organ lesions also abolish drinking responses to systemic injections of angiotensin.

There is, however, some evidence that the septal region may exert significant and specific influences on angiotensin-related thirst. Septal lesions result in persistent hyperdipsia (Harvey & Hunt, 1965), and Lubar, Boyce, and Schaefer (1968) have provided evidence that this effect of the lesion is not secondary to polyuria. Blass and Hanson (1970) have reported that rats with septal lesions responded normally to hypertonic saline challenges but drank more than controls when given subcutaneous polyethylene glycol injections. This suggested a selective facilitation of the response to extracellular hypovolemia, but a series of recent papers (Black & Mogenson, 1973; Blass, Nussbaum, & Hanson, 1974) have indicated that not all treatments that result in extracellular hypovolemia are affected by septal lesions. Clear facilitatory effects were seen only when drinking was induced by renin or angiotensin or by treatments (such as caval ligation, hypotension due to isoproterenol, etc.) which affect the renin-angiotensin system.

It is clear from this brief and necessarily selective overview of some recent developments in this field that much remains to be learned about the functions of various extrahypothalamic structures in the regulation of food and water intake. It is already clear, however, that a number of subcortical regions outside the hypothalamus may exert many regulatory influences which were assigned to hypothalamic mechanisms only a few years ago. The emerging pattern suggests that we must take a closer and more critical look at the data that have led to the postulation of hypothalamic control.

ARE FIBERS OF PASSAGE RESPONSIBLE FOR THE EFFECTS OF LATERAL HYPOTHALAMIC LESIONS?

The principal experimental support for a hypothalamic model of food and water intake regulation comes from studies that have shown severe impairments after lesions in the dorsolateral quadrant of the hypothalamus. This area is traversed by numerous diffuse fiber systems (e.g., components of the medial forebrain bundle, the

pallidofugal fiber system, nigrostriatal pathway, various corticofugal fiber systems, and important sensory projections). The sizable lesions, which result in aphagia and adipsia, interrupt many of these fiber systems and may thus interfere with feeding, but not because of the destruction of cellular components of the lesion site. It has also become increasingly clear that lateral hypothalamic lesions produce sensory and motor disturbances that may themselves contribute importantly to the impaired feeding and drinking behavior.

That the effects of hypothalamic lesions might be due to an interruption of fibers of passage rather than the destruction of cellular components of the area was first suggested by Morgane (1961a), when he found that lesions in the globus pallidus, which interrupt the pallidofugal projections that pass through the lateral hypothalamus on their way to the lower brainstem, reproduced the effects of lateral hypothalamic lesions. Gold (1967) supported this interpretation by showing that lesions in the brainstem tegmentum, which is one of the principal targets of the pallidofugal projection system, also produced aphagia and adipsia, and that a unilateral lesion in that area, when added to a contralateral lateral hypothalamic lesion, also produced persistent aphagia and adipsia. Teitelbaum and Krantz's (cited in Teitelbaum, 1967) observation of similar additive interactions between unilateral hypothalamic and pallidal lesions further strengthens the suspicion that the effects of lateral hypothalamic lesions may be due to an interruption of the pallidofugal fiber system, which has no known synapses in the area of the feeding center.

Ungerstedt (1971) has more recently implicated the nigrostriatal pathway, which ascends through the lateral hypothalamus and adjacent portions of the internal capsule without known synapse. He reported that electrolytic lesions of the substantia nigra, or microinjections of 6-hydroxydopamine (6-OHDA), which selectively destroy the catecholaminergic components of the region, produce persistent aphagia and adipsia (as well as akinesia

and reduced alertness). Ungerstedt's hypothesis that the nigrostriatal pathway may be responsible for the effects of lateral hypothalamic lesions on food and water intake is not incompatible with the earlier observation that lesions in the globus pallidus result in aphagia and adipsia, because major components of the nigrostriatal pathway course through the globus pallidus before terminating in the caudate and putamen (Ungerstedt, 1971). Lesions in the globus pallidus (Andén, Fuxe, Hamberger, & Hökfelt, 1966) and the lateral hypothalamus (Glick, Greenstein, & Waters, 1974; Olmians & Harvey, 1972) deplete striatal dopamine, proof positive that this pathway is, indeed, interrupted in several of the lesions that produce aphagia and adipsia.

Ungerstedt's observations have been replicated and related not only to the typically transient aphagia and adipsia but also to persisting regulatory deficits in food and water intake. Fibiger, Zis, and McGeer (1973) have reported that injections of 6-OHDA into the substantia nigra produce a transient (approximately 10 days) aphagia and adipsia. Following recovery of voluntary ingestive behavior, the animals showed many of the persisting deficits that characterize rats with lateral hypothalamic lesions. Drinking occurred mainly in association with food intake, finickiness was increased, the response to experimentally induced intracellular dehydration was absent, and the anorexigenic effects of amphetamine were significantly reduced. Marshall and his associates (Marshall, Richardson, & Teitelbaum, 1974; Marshall & Teitelbaum, 1973) have observed more persistent aphagia and adipsia after injection of 6-OHDA into the substantia nigra or into the globus pallidus. These animals showed the gradual recovery of voluntary intake that Teitelbaum and Epstein (1962) observed in rats with lateral hypothalamic lesions. They subsequently displayed persisting deficits in their response to cellular glucoprivation, food deprivation, which were essentially indistinguishable from those seen after

electrolytic lesions of the lateral hypothalamus (the 6-OHDA injections produced somewhat less finickiness). The results of these experiments do not, unfortunately, permit an unqualified conclusion concerning the role of the dopaminergic nigrostriatal fiber system in food and water intake regulation, because the injection of 6-OHDA into the substantia nigra, as well as the globus pallidus, produced marked depletions of forebrain and striatal norepinephrine (as well as sizable depletions of dopamine). The duration of the aphagia correlated well with the extent of dopamine depletion in the striatum, but the severity of the persisting regulatory deficits, surprisingly, did not. (A similar correlation between the duration of aphagia and adipsia and the degree of striatal dopamine depletion after electrolytic lesions in the lateral hypothalamus has been reported by Oltmans & Harvey, 1972). Smith, Strohmayer, and Reis (1972) have reported that injections of 6-OHDA directly into the lateral hypothalamus produced aphagia and adipsia, although the effects were not as persistent as those often seen after lateral hypothalamic lesions. Their severity and duration appeared to be correlated with the effectiveness of the injections in reducing hypothalamic norepinephrine rather than striatal dopamine.

These investigations demonstrate that noradrenergic components of the hypothalamus, striatum, and telencephalon are affected by 6-OHDA injections into the lateral hypothalamus, as well as into neighboring tissues such as the substantia nigra, the globus pallidus, or the ventricles. This is a potentially important problem, because noradrenergic mechanisms have been implicated in the regulation of feeding behavior in other experimental paradigms. Perhaps the most substantial body of evidence suggesting that central noradrenergic pathways participate in the organization of feeding-related behaviors comes from the investigations that have shown that microinjections of adrenergic compounds and related drugs into the diencephalon elicit or inhibit food intake selectively. The initial experiments (Gross-

man, 1960, 1962a, 1962b) demonstrated that the application of norepinephrine to the perifornical region of the hypothalamus elicited eating in sated rats, and that injections of compounds that interfere with the action of norepinephrine blocked this effect and inhibited feeding in deprived animals. Subsequent investigations of this phenomenon (e.g., Booth, 1967; Coury, 1967; Grossman, 1964a, 1964b; Leibowitz, 1970) have shown that feeding can be elicited by injections of alpha adrenergic compounds into various other parts of the hypothalamus, as well as into sites anterior and dorsal to it. The doses used in the initial investigations were quite high, so that the possibility of a nonspecific drug action could not be excluded. However, more recent studies (e.g., Ritter & Epstein, Note 2; Leibowitz, Note 3) have succeeded in reducing the dosage to more nearly physiological levels, and the conclusion that feeding may be influenced by noradrenergic components of the upper brainstem appears well supported at this time.

Reports that intraventricular injections of 6-hydroxydopa (which deplete brain norepinephrine but not dopamine) produce aphagia (Richardson & Jacobowitz, 1972) and that intraventricular injections of norepinephrine reverse the aphagia seen in rats after lateral hypothalamic lesions (Berger, Wise & Stein, 1971) are also in agreement with such an interpretation.

The results of a number of recent investigations do, however, suggest that the destruction of dopaminergic pathways may be responsible for many of the disturbances in food and water intake that are seen after intracerebral injections of 6-OHDA. Ungerstedt (Note 4) has recently reported that the administration of dopamine agonists elicits feeding in rats that are aphagic as a result of 6-OHDA injections into the substantia nigra. Zigmond and Stricker (1972) reported earlier that intraventricular injections of this compound depleted brain norepinephrine by 95% (but dopamine by only 62%) and produced only mild and transient hypophagia. When the animals were pretreated with the

monoamine oxydase inhibitor pargyline, no further depletion of norepinephrine was observed, but dopamine was reduced to 5% of normal and more persistent aphagia and adipsia occurred.

In a recent extension of this investigation, Stricker and Zigmond (1974) observed that rats given intraventricular injections of 6-OHDA plus a monoamine oxydase inhibitor progressed through the gradual and orderly recovery stages described by Teitelbaum and Epstein (1962) for rats with lateral hypothalamic lesions. After recovery of voluntary intake, these animals failed to eat in response to cellular glucoprivation and showed only the delayed drinking response to extracellular hypovolemia that is typically seen in rats with lateral hypothalamic lesions. However, these animals responded promptly to hypertonic saline challenges and drank normal quantities of water during periods of food deprivation. This apparent sparing of at least some of the mechanisms that regulate water intake is surprising in view of recent reports of persisting deficits after 6-OHDA injections into the substantia nigra or globus pallidus (Fibiger et al., 1973; Marshall et al., 1974).

The effects of intraventricular injections of 6-OHDA support the hypothesis that at least some components of the lateral hypothalamic lesion syndrome may be related to an interruption of dopaminergic fibers of passage. They also raise some interesting questions and suggest that this may not be the entire story. One may ask, for instance, why voluntary food and water intake return even though the dopamine levels of the striatum remain near zero levels in these experiments, or why lateral hypothalamic lesions, which typically result in much less complete depletions of striatal dopamine (Glick et al., 1974; Olmians & Harvey, 1972), have more severe and persisting effects on food and water intake. In this context, one also must not lose sight of the fact that aphagia and adipsia have been reported after lesions in portions of the brainstem that are not traversed by the dopaminergic nigrostriatal system (e.g., Lyon, Halpern, & Mintz, 1968; Parker &

Feldman, 1967; Zeigler & Karten, 1974). Although it is possible that these lesions may interfere with afferents to the substantia nigra, it is unlikely that they could result in the massive striatal dopamine depletions that are apparently necessary to produce persistent changes in food and water intake. The effects of these brainstem lesions on hypothalamic, striatal, or telencephalic norepinephrine are, unfortunately, not known at this time.

These considerations suggest that the effects of lateral hypothalamic lesions on food and water intake may not be entirely determined by the destruction of a single regulatory mechanism but may, instead, reflect a complex syndrome of disturbances in several functions, which may be mediated by several distinct fiber systems.

We have begun to investigate the functions of the fiber systems that interconnect the brainstem and basal ganglia in some detail (Grossman & Grossman, 1971, 1973). We find that surgical transections of all fibers that enter or leave the hypothalamus laterally (including pallidofugal and nigrostriatal pathways and the reciprocal connections between the brainstem and temporal lobe) reproduce the effects of lateral hypothalamic lesions in every respect. In these experiments, a retractable tungsten-wire knife of 125- μ m diameter was used to make a cut in the parasagittal plane along the medial border of the internal capsule. In the rat, a cut that is 3–4 mm in length and extends about 2 mm above the bottom of the brain effectively severs most of the lateral connections of the hypothalamus without producing significant direct damage to cellular components of the hypothalamus or striatum. We find that such a cut produces long-term aphagia and adipsia which in many instances cannot be reversed within 90–120 days, in spite of concerted efforts to wean the animals to a highly palatable diet. Animals that recover voluntary food and water intake undergo the gradual recovery described by Teitelbaum and Epstein (1962) for rats with lateral hypothalamic lesions. In the last stage of this recovery process, our rats consume adequate amounts of a palatable,

dry, pellet diet and are, indeed, hyperphagic when a preferred liquid diet is offered. However, they promptly return to a state of hypophagia or aphagia when a less preferred, dry, powdered diet (which normal rats eat readily) is presented or when quinine is added to a palatable liquid diet. This pattern of ingestive behavior suggests that these recovered animals are extremely sensitive to the sensory qualities of their diet, and there is reason to believe that palatability may be the principal and perhaps sole determinant of food ingestion in these animals.

The importance of central glucoreceptor mechanism has recently been questioned by the observation that large anterior hypothalamic lesions (Blass & Kraly, 1974), as well as damage to the zona incerta (Walsh & Grossman, unpublished observation, 1974) or ventromedial mesencephalon (Alheid & Grossman, unpublished observation, 1974), abolish the feeding response to cellular glucoprivation without resulting in significant disturbances in ad-libitum food intake. It is nonetheless commonly believed that blood sugar or, more precisely, its utilization by central glucoreceptors may be an important signal for the regulation of food intake. This hypothesis is based in part on the observation that systemic injections of insulin, which lower blood sugar and presumably its intracellular utilization, increase food intake (Mackay, Calloway, & Barnes, 1940), and that systemic administration of the glucose analogue 2-deoxy-D-glucose, which competes with glucose for intracellular metabolic pathways, produces similar effects (Smith & Epstein, 1969). It has been shown that rats with lateral hypothalamic lesions do not respond to these treatments (Epstein & Teitelbaum, 1967; Wayner et al., 1971). We similarly find that rats that recovered from the effects of transections of the lateral connections of the hypothalamus did not eat in response to insulin or 2-deoxy-D-glucose (Grossman & Grossman, 1973). Indeed, our animals also failed to adjust food intake reliably in response to changes in environmental temperatures—the only regulatory response

which appears to be intact after lateral hypothalamic lesions (Epstein & Teitelbaum, 1967)—or injection of 6-OHDA into the substantia nigra or globus pallidus (Marshall et al., 1974; Marshall & Teitelbaum, 1973).

An examination of the water intake of our rats following apparent recovery of voluntary intake suggests that they do not regain any regulatory capabilities following surgery but drink water only to facilitate the ingestion of dry food. Water intake falls to zero during periods of food deprivation, and there is no response to treatments that elicit or increase drinking in the normal rat.

Drinking occurs mainly in response to the development of intracellular dehydration or to a decrease in the volume of extracellular fluid. The operation of the first mechanism is readily demonstrated in the intact rat by the subcutaneous or intraperitoneal administration of hypertonic saline, which results in cellular dehydration and elicits water intake (e.g., Corbit, 1965). Extracellular hypovolemia can be experimentally produced by subcutaneous or intraperitoneal injections of hyperoncotic compounds, such as polyethylene glycol, which sequester extracellular fluids and result in intense thirst, which is not slaked by water (Stricker, 1968). The two mechanisms become effective at different times during ontogeny (Friedman & Campbell, 1974) and are believed to be independent. Drinking is also induced by the peripheral or central administration of renin, or its metabolic product angiotensin, and by a variety of pharmacological or physiological treatments that induce the release of this hormone from the kidney. There is also increasing evidence that hypotensive agents such as isoproterenol elicit thirst by acting on still a fourth and possibly independent mechanism (see recent reviews by Epstein, 1973; Fisher, 1973; and Lehr, 1973).

Rats with lateral hypothalamic lesions do not drink in response to extracellular hypovolemia or intracellular dehydration, although long-term adjustments in intake have been reported (see Stricker, 1973,

for review), and a similar loss of regulatory control was seen in our rats that had recovered from the effects of knife cuts along the border of the lateral hypothalamus (Grossman & Grossman, 1973).

Our observations indicate that an interruption of the fibers that enter or leave the lateral border of the hypothalamus reproduce, in detail, the effects of lateral hypothalamic lesions on food and water intake. This lends support to the hypothesis that possibly all of the effects of lateral hypothalamic damage on ingestive behavior may be due to an interruption of fiber systems that may or may not synapse in the area identified as the lateral hypothalamic "feeding region." Since this area is traversed by several major fiber systems that interdigitate extensively (see Dahlström & Fuxe, 1964; Guillery, 1957; Knook, 1965; Millhouse, 1969, 1973a, 1973b; Nauta, 1957, 1958; Raisman, 1966; Ungerstedt, 1971; Wolf & Sutin, 1966), it has been difficult to ascertain which of them might be responsible for the observed effects on food and water intake. We have made some progress in this respect, but much remains to be learned.

Morgane (1961b, 1961c) reported a number of years ago that lesions that interrupted the medial forebrain bundle anterior or posterior to the region of the lateral hypothalamic feeding center did not produce lasting aphagia or adipsia. Similar results have been reported by DiCara and Wolf (1968). We have found more recently that small knife cuts in the coronal plane, just rostral to the substantia nigra, produced long-term (>30 days) aphagia and adipsia (Kent, Rezak, & Grossman, 1973). A similar cut only .5 mm rostral to this plane and even quite large cuts across the medial forebrain bundle, which severed most or all of its lateral hypothalamic components just behind the feeding center, had only temporary (1-6 days) inhibitory effects on food and water intake (McDermott & Grossman, unpublished observation, 1973). Paxinos and Bindra (1973) also reported normal food intake following similar coronal cuts across the medial forebrain bundle.

Albert, Storlien, Wood, and Ehman (1970) have reported aphagia and adipsia after very large coronal knife cuts anterior and posterior to the lateral hypothalamic feeding center, but it appears likely that their results were due to an involvement of fibers that travel in the internal capsule. The overall pattern of results seen after knife cuts across the medial forebrain bundle suggests that fibers that are related to ingestive behavior may not ascend very far in the lateral hypothalamus itself, but turn laterally and perhaps dorsally in the posterior diencephalon. Such an interpretation is supported by the results of an earlier investigation (Blatt & Lyon, 1968), which demonstrate additive effects on food and water intake of asymmetric lesions in the midbrain and in lateral portions of the subthalamic region and parts of the zona incerta.

We have found that parasagittal cuts in the striatum and internal capsule, which run parallel to those that result in long-term aphagia and adipsia, do not produce these effects (Grossman & Grossman, 1971; Grossman & Grossman, unpublished observation, 1973). Alheid and I (Alheid & Grossman, unpublished observation, 1974) have further observed that short parasagittal cuts along the lateral border of the preoptic area and anterior hypothalamus do not produce persisting aphagia or adipsia, even though they are in the same parasagittal plane as more posterior cuts, which are consistently effective. Negative results were also observed in rats with lesions in each of the major nuclei of the amygdala and in the piriform cortex, which, in the rat, is probably the sole source of the amygdalofugal pathway (Grossman, Grossman, & Walsh, 1975; Lammers, 1972). These results suggest that reciprocal connections between the brainstem and temporal lobe, including lateral projections of the medial forebrain bundle, the amygdalofugal pathway, and ansa peduncularis of Meynert, are probably not responsible for the effects of our cuts along the lateral border of the hypothalamus. Paxinos (1974) has recently reported results which support this con-

clusion. It is possible that the critical pathways may turn rostrally immediately after entering the internal capsule and project to or from cortical sites (Webster, 1961, 1965). However, it appears likely, on the basis of the work of Morgane (1961a, 1961b, 1961c), Gold (1967), Ungerstedt (1971), Marshall et al. (1974), and others, that components of the complex and diffuse feedback system that interconnects the brainstem and striatum may be involved.

DO SENSORY AND MOTOR DYSFUNCTIONS PLAY AN IMPORTANT PART IN THE LATERAL HYPOTHALAMIC SYNDROME?

The pallidofugal and nigrostriatal projections play an important role in the feedback system that controls posture, locomotion, and other movements (see Laursen, 1963, for review). The effects of our knife cuts, as well as those of lateral hypothalamic, pallidal, and striatal lesions, might, therefore, reflect motor or sensory disturbances rather than an interference with mechanisms specifically related to hunger or thirst. Rats with lateral hypothalamic lesions are not only aphagic and adipsic but also are extremely hypoactive—although they display starvation hyperactivity after prolonged deprivation (Campbell & Baez, 1974). They often display symptoms similar to those seen in catatonic patients (e.g., they remain motionless in very unusual postures for long periods of time), fail to groom themselves, and do not respond promptly and efficiently to sensory stimuli (Balagura, Wilcox, & Coscina, 1969; Marshall, Turner, & Teitelbaum, 1971; Rodgers, Epstein, & Teitelbaum, 1965; Teitelbaum & Epstein, 1962). Many of these symptoms disappear or diminish while the animals are still aphagic and adipsic, and there is some evidence that the duration of the aphagia does not correlate with the severity or persistence of the overt motor dysfunctions (Balagura et al., 1969). This does not, however, rule out more subtle impairments in the muscular control necessary to chew, bite, lick, swallow, etc., or the "sensory-neglect,"

which Marshall and his associates (Marshall et al., 1971; Marshall & Teitelbaum, 1974) have seen in association with aphagia and adipsia. Rodgers et al. (1965) have shown that rats with lateral hypothalamic lesions not only fail to feed and drink but also do not press a lever for intragastrically delivered food rewards. This observation supports a motivational interpretation, but it is not as conclusive as one might wish, because the animals' ability to lever-press for other reinforcements (such as brain-stimulation rewards or shock avoidance) was not demonstrated. That this is an essential control is suggested by several reports of impaired brain-stimulation-rewarded or food-rewarded lever pressing in animals with lesions just rostral or caudal to the feeding center (Chase & Moore, 1968; DiCara & Wolf, 1968; Morgane, 1961b, 1961c). Severe impairments of escape and avoidance behaviors (Coscina & Balagura, 1970; Runnels & Thompson, 1969) as well as food-rewarded or brain-stimulation-rewarded lever pressing (Morgane, 1961b, 1961c) have been reported after lesions in the lateral hypothalamic feeding center itself. The problem is emphasized by our own observation (Kent & Grossman, 1973; Kent et al., 1973; Alheid & Grossman, unpublished observation, 1973) that many of the knife cuts which duplicate the effects of lateral hypothalamic lesions on ingestive behavior also abolish lever pressing (as well as other instrumental behaviors), irrespective of reward conditions.

Investigators who have produced aphagia and adipsia by electrolytic or chemical lesions outside the hypothalamus have frequently commented on the severity of the accompanying motor disturbances and have often suggested that the disturbances in feeding behavior might not reflect motivational impairments. Morgane (1961b), for instance, suggested that lesions in the globus pallidus might result in basic metabolic dysfunctions; Blatt and Lyon (1968) proposed that the aphagia and adipsia seen after damage to the midbrain tegmentum might be due to an interruption of pathways which "normally aid in the

association of sensory cues and motor actions necessary for the finding, seizing, and introducing into the mouth of food objects" (p. 576); Ungerstedt (1971) noted a marked akinesia and lack of alertness in rats after injections of 6-OHDA into the substantia nigra, and discussed the possibility that this might be the cause of the observed aphagia and adipsia; and Schwartzbaum and his associates (Levine, Ferguson, Kreinick, Gustafson, & Schwartzbaum, 1971; Levine & Schwartzbaum, 1973) have made similar comments concerning the effects of striatal lesions on ingestive behavior, following a thorough study of the sensorimotor capabilities of rats with unilateral or bilateral striatal damage. Zeigler and Karten (1974) recently reported aphagia and adipsia after lesions in the brainstem and basal thalamus and suggested that these effects might be directly related to an interruption of sensory signals carried by the trigeminal nerve from the region of the mouth.

Our own observations of rats that are aphagic and adipsic as a result of knife cuts lateral to or behind the feeding center do not, as yet, permit safe conclusions concerning the possible role of sensory or motor disturbances. We typically observe motor dysfunctions that are much less pronounced and persistent than those produced by most lateral hypothalamic lesions. Hypoactivity and catatonia-related symptoms are generally mild and disappear within a week (whereas aphagia has been seen for many months), grooming is normal, and there is little evidence of "sensory-neglect" as described by Marshall et al. (1971). However, after recovery of voluntary ingestive behavior, our animals have apparent difficulties with gnawing, biting, chewing, and perhaps swallowing (as indicated by an increase in the amount of shredded waste under their cages), and we have observed some awkwardness in licking food or water from a tube, particularly when its tip did not protrude into the animal's cage but was slightly recessed. To what extent these motor dysfunctions may contribute to the aphagia and adipsia syndrome is, at the moment, not clear.

It is often argued (e.g., Epstein, 1971) that the persisting regulatory deficits, which are seen in rats with lateral hypothalamic lesions (as well as in our rats with cuts lateral to the lateral hypothalamus and in rats with nigral or pallidal damage) after recovery of voluntary intake, are not amenable to such an interpretation, because the animals are demonstrably capable of the act of feeding and drinking. It is, however, possible that the presence of residual motor or sensory deficits may make ingestive behavior more difficult and may thus increase the threshold of eating or drinking in response to insulin, 2-deoxy-D-glucose, and extracellular or intracellular dehydration, which may not be as powerful as is generally assumed (the response to insulin and 2-deoxy-D-glucose, in particular, is often small and variable even in controls). The fact that recovered animals almost always drink normal quantities of water in the presence of food but are adipsic when food deprived suggests that the adipsia may be related to a dysfunction of a motivational nature, but a possible contribution of sensorimotor impairments at some stage of the recovery cycle clearly cannot be ruled out. Clarification of this question and of the role of sensory and motor dysfunctions in the typically transient aphagia is, without doubt, one of the most important tasks in this field.

POSSIBLE CONTRIBUTION OF FIBER SYSTEMS TO THE EFFECTS OF VENTROMEDIAL HYPOTHALAMIC LESIONS

Although hyperphagia-inducing hypothalamic lesions typically involve the ventromedial nucleus, it has long been known that the most effective lesions are located lateral and slightly ventral to the nucleus itself (Brobeck, 1946; Graff & Stellar, 1962; Hetherington & Ranson, 1942), and that lesions that are restricted to the nucleus itself do not produce hyperphagia (Joseph & Knigge, 1968). Since many of the effective lesions are indeed caudal to the ventromedial nucleus (Graff & Stellar, 1962; Hetherington & Ranson, 1942), it was suggested many years ago that the hyperphagia might be due to

destruction of projections to or from the lower brainstem (Brobeck, 1946). Gold (1973) has recently re-examined the effects of lesions in the ventromedial hypothalamus in the light of new evidence implicating the ventral noradrenergic pathway in feeding and satiety (discussed later), and concluded from his investigation that only lesions which interrupt this bundle ventrolateral to the ventromedial nucleus result in the classic hyperphagia syndrome.

The ventral noradrenergic bundle arises from cell groups in the medulla and pons and ascends through the ventral tegmentum. Its axons innervate medial portions of the diencephalon preferentially, although longer projections to other limbic structures have been observed (Fuxe, 1965; Fuxe & Hökfelt, 1969). It is of some interest in the context of our discussion that only a few noradrenergic axons terminate on cells of the ventromedial nucleus itself and that the system projects profusely to the paraventricular nucleus (Fuxe, 1965), which appears to be a focal point for feeding elicited by microinjections of norepinephrine into the hypothalamus (Leibowitz, Note 3). The ventral noradrenergic bundle has also been implicated by Ahlskog and Hoebel's (1973) report of hyperphagia after electrolytic or chemical lesions in the ventral tegmental region that is traversed by this bundle, which depleted diencephalic (as well as telencephalic) norepinephrine.

That the effects of ventromedial hypothalamic lesions on food intake and body weight may be due to an interruption of the ventral noradrenergic bundle is also indicated by our observation (Grossman, 1971) that knife cuts in the coronal plane caudal to the ventromedial nucleus (which spare the tissues that are most commonly involved in hyperphagia-inducing lesions, but interrupt the posterior connections of the medial hypothalamus, including the ventral noradrenergic bundle) also produced marked hyperphagia. Hennessy and I (unpublished observation, 1974) have recently studied the effects of these cuts in more detail, because Sclafani (1971), working in my laboratory, did not observe significant changes in food intake or body

weight following knife cuts immediately behind the tip of the ventromedial nucleus. Our investigation disclosed that only cuts which transected fibers coursing ventral and lateral to the region behind the ventromedial nucleus consistently produced hyperphagia. This observation is congruent with the hypothesis that a transection of components of the ventral noradrenergic bundle may be responsible for this effect. We also observed that the effects of such cuts on food intake and body weight are prominent and persistent only in female rats. Males showed marked hyperdipsia, but only more transient and relatively mild effects on food intake, which might easily escape detection in experiments that use male subjects exclusively, such as those reported by Paxinos and Bindra (1972).

We (Sclafani, 1971; Sclafani & Grossman, 1969) as well as a number of other investigators (e.g., Albert & Storlien, 1969; Gold, 1970) have also observed hyperphagia following parasagittal knife cuts between the medial and lateral hypothalamus. It was originally assumed that this might be due to an interruption of direct interconnections between the ventromedial region and adjacent parts of the lateral hypothalamus, which recent neuroanatomical investigations had demonstrated (Arees & Mayer, 1967; Millhouse, 1973a, 1973b). However, a more detailed analysis of the histological materials (Gold, 1970; Sclafani, 1971) has shown that the most effective cuts are anterior and dorsal to most of the ventromedial nucleus and anterior to the lateral feeding area itself. Gold, Quackenbush, and Kapatos (1972) have reported that a unilateral cut in this region (which produced only minor and transient changes in food intake) can be combined with a contralateral lesion in the region of the mammillary bodies or in the ventral tegmentum to produce persisting hyperphagia. The possibility of a hypothalamic synapse is certainly not excluded by these observations, but they suggest that the same fiber system may be involved in hyperphagia-producing lesions and coronal cuts posterior to the ventromedial hypo-

thalamus and in the parasagittal cut dorso-lateral to it (as well as in lesions in the area of the ventromedial hypothalamus itself).

A review of the effects of coronal cuts through the anterior hypothalamus suggests that the dorsolateral projections to or from the ventromedial hypothalamus, which are responsible for the hyperphagia in these experiments, may ascend for some distance in the medial segment of the anterolateral hypothalamus. Sclafani (1971), working in my laboratory, as well as others (Palka, Coyer, & Critchlow, 1969; Storlien & Albert, 1972) have found that small coronal cuts immediately rostral to the tip of the ventromedial nucleus do not significantly modify food intake or body weight. I have observed transient hyperphagia and exaggerated weight gain after larger cuts in the same coronal plane, which extended into the medial aspects of the lateral hypothalamus (Grossman, 1971). Other investigations have subsequently reported moderate hyperphagia and increased weight gain in male (Paxinos & Bindra, 1972) as well as female (Storlien & Albert, 1972) rats, after extensive coronal cuts through the anteromedial hypothalamus.

Hyperphagia has also been reported following knife cuts in the horizontal plane, which effectively isolate the ventromedial nucleus (and overlying regions) from the paraventricular nucleus (Palka, Liebelt, & Critchlow, 1971). These observations are compatible with the hypothesis that an interruption of components of the ventral noradrenergic bundle may contribute to the hyperphagia that is seen after various hypothalamic lesions. They may be particularly significant in view of Leibowitz' (Note 3) recent report that the paraventricular nucleus may be a focus for feeding elicited by microinjections of noradrenergic compounds.

Although most experimental observations are compatible with this interpretation, one should keep in mind that the suggested relationship between satiety and noradrenergic components of the medial hypothalamus is based almost entirely on inferences from possibly misleading corre-

lations. There is a general agreement that the ventral noradrenergic bundle arises in the midbrain and projects mainly to many portions of the diencephalon, but its exact course and distribution are not yet well known. Moreover, the noradrenergic projections are diffuse and overlap extensively, with a wide variety of other neural pathways that project to the medial hypothalamus, and it is by no means clear that we are justified in focusing our attention on the monoaminergic projections.

We have only recently begun to understand the anatomical organization of the ventromedial hypothalamus in any detail, and it is not yet clear how the behavioral effects of lesions and knife cuts in and around the area are related to the anatomical data. Until very recently it was believed that only the anterior hypothalamus provided significant afferent projections to the ventromedial nucleus (Chi, 1970), and it was widely held that the nucleus had very few if any efferent projections beyond its border (Kaelber & Leeson, 1967; Sutin & Eager, 1969), leaving it "enveloped in splendid isolation," as one reviewer recently put it (Millhouse, 1973a). Only within the last few years has it been demonstrated that the outer shell of the nucleus (a) receives afferents from the stria terminalis, the mammillary peduncle, the medial forebrain bundle, and the lateral hypothalamus; (b) projects to the lateral hypothalamus, the anterior hypothalamus, the zona incerta, and the dorsolateral hypothalamus; and (c) sends fibers caudally into the periventricular fiber system and medial forebrain bundle (Eager, Chi, & Wolf, 1971; Millhouse, 1973a, 1973b).

HORMONAL INFLUENCES ON HYPOTHALAMIC HYPERPHAGIA

The role of steroid hormones in the etiology of hypothalamic hyperphagia has been the subject of some debate. There is no doubt that sex hormones influence food intake and body weight, and many other observations suggest that their effects may be mediated by hypothalamic mechanisms. The nature of this relationship is as yet poorly understood.

In the male rat, castration decreases food intake, but increases body weight because of a sharp decrease in activity (Kakolewski, Cox, & Valenstein, 1968). Injections of testosterone propionate return food intake to normal levels (Hervey & Hutchinson, 1973). In the female rat, food intake and activity covary with the estrous cycle (Brobeck, Wheatland, & Strominger, 1947; Kennedy & Mitra, 1963); ovariectomy increases food intake (Kakolewski et al., 1968); and injections of estradiol benzoate reverse this effect (Zucker, 1969, 1972). It is particularly interesting that estradiol does not modify food intake unless the animal is at an abnormally high body weight (Redick, Nussbaum, & Mook, 1973; Zucker, 1972).

These and related observations suggest that the steroid hormones may act on the brain to alter the "set-point" at which body weight is regulated (e.g., Mook, Kenney, Roberts, Nussbaum, & Rodier, 1972; Wade, 1972). Similar set-point hypotheses have been advanced to describe the role of the ventromedial (Hoebel & Teitelbaum, 1966) or lateral (Powley & Keesey, 1970) hypothalamus, and there is considerable evidence that steroids may be acting on hypothalamic mechanisms (although not exclusively so) in order to modify food intake.

Lesions in the ventromedial hypothalamus (Kennedy, 1957) do not produce hyperphagia in immature female rats until they reach puberty. Although prepuberal rats (like ovariectomized adults) are hyperphagic when food intake is considered in relation to body weight, the ineffectiveness of the hypothalamic lesion does not appear to reflect simple ceiling effects, because prepuberal rats are quite capable of increasing their food intake when exposed to cold (Teitelbaum, Cheng, & Rozin, 1969). In the adult marked sex differences have been reported, but the generality of the phenomenon is uncertain. Valenstein and his associates (Cox, Kakolewski, & Valenstein, 1969; Valenstein, Cox, & Kakolewski, 1969) have reviewed numerous reports of ventromedial hyperphagia and concluded that the effects of ventromedial hypo-

thalamic lesions on food intake and body weight were significantly larger and more readily obtainable in the female. A direct experimental test provided support for this hypothesis. The results of more recent investigations (e.g., Gold, 1970; Rehovsky & Wampler, 1972) indicate that this sex difference may not be as general a phenomenon as the earlier reports suggested, but the possibility of hormonal influences has not been ruled out.

Ventromedial hypothalamic lesions disrupt female mating behavior (Kennedy & Mitra, 1963), and there is some controversy concerning their effects on gonadal function. Valenstein et al. (1969) reported evidence of gonadal atrophy after ventromedial lesions, but Wampler and Gier (1974) recently presented evidence suggesting that male as well as female gonads retain their capability to secrete steroid hormones after ventromedial hypothalamic lesions and may, in fact, be unaffected by conditions (such as constant illumination) that interfere with cyclic hormone release in the female rat. Ventromedial hypothalamic lesions increase the release of prolactin, which has been shown to increase food intake in the female, but not in the male (Pfaff, 1969). The ventromedial hypothalamus is most directly implicated by Wade and Zucker's (1970) report that implantation of estradiol into that area, but not into adjacent regions, decreased food intake within 12 hours in spayed female rats. Similar observations were recently reported by Jankowiak and Stern (1974).

Direct comparisons between the effects of ovariectomy and ventromedial hypothalamic lesions indicate some remarkably detailed similarities (e.g., both increase meal size in the light as well as in the dark, but increase meal frequency in the light only [Balagura & Devenport, 1970; Kenney & Mook, 1974]) but also some clear differences (e.g., lesions increase finickiness, while ovariectomy has opposite effects [Teitelbaum, 1955; Zucker, 1969]).

These observations are congruent with the hypothesis that estradiol may affect food intake by acting on components of

the ventromedial hypothalamus. However, others indicate that there must be other sites of action and that the effects of lesions in this area cannot be entirely due to a loss of normal steroid influences on food intake. The first of these reservations is suggested by the fact that ovariectomy increases, and subsequent hormone treatments decrease, food intake in rats that are hyperphagic as a result of ventromedial hypothalamic lesions (King & Cox, 1973; Montemurro, 1971). Nor does the critical site of action seem to lie in projections to the lateral hypothalamus, as one might conclude from the observation that females show a permanently lowered body weight set-point after lateral hypothalamic lesions (Powley & Keesey, 1970), whereas males do not (Cox & Kakolewski, 1970). This is indicated by the fact that ovaries exercise a restraining effect on body weight in rats that have recovered from the effects of lateral hypothalamic lesions (Cox & Kakolewski, 1970; Harrell & Balagura, 1975). The second point—that lesions in the ventromedial hypothalamus do not merely prevent the action of steroid hormones—is indicated by the fact that the effects of lesions are much more severe and persistent than those seen after ovariectomy. A 100% increase in body weight is commonly seen after lesions, whereas ovariectomy typically results in an increase of only 20–25% (Mook et al., 1972). The argument against such an interpretation is further strengthened by the fact that the hyperphagia that results from ventromedial hypothalamic lesions is apparently not prevented by hypophysectomy (Cox et al., 1969) but is markedly affected by subdiaphragmatic vagotomy (Powley & Opsahl, 1974).

It is clear that we do not yet understand the etiology of the hypothalamic hyperphagia, but it appears likely that it may have several independent components. Some observations from my own laboratory, as well as others, suggest that some of the contributing factors can be isolated by recently developed neuroanatomical or neurochemical techniques.

Hennessy and I (unpublished observation, 1974) have recently completed a detailed investigation of the effects of coronal knife cuts in the posterior hypothalamus and have found that the hyperphagia that reliably occurs after some of these cuts differs in some critical respects from that seen after electrolytic lesions in the ventromedial hypothalamus. Female rats displayed a pronounced hyperphagia, but the effect was relatively short-lived (5–8 weeks) and rarely resulted in a total increase in body weight in excess of about 25%. Most male rats overate only slightly, if at all, shortly after surgery but displayed a dramatic hyperdipsia (the animals consumed 200–300 ml per day) that persisted unabated for many months. These animals had renal dysfunctions similar to those seen in diabetes insipidus (even though the knife cuts did not involve the infundibular stem), suggesting that the excessive water intake may have been related to an interference with antidiuretic hormone release by the pituitary (Grossman, Hennessy, & Kanner, unpublished observation, 1974). In many animals the excessive water intake was, however, entirely feeding related, and water consumption dropped to normal or even below normal levels during periods of food deprivation, suggesting that other factors may have contributed to the hyperdipsia. Kent and I (unpublished observation, 1974) have recently observed hyperdipsia in animals with coronal cuts in the ventromedial mesencephalon, several millimeters behind the infundibulum. We have not yet tested kidney functions in these animals, but it appears unlikely that this cut would interfere with pituitary hormone release. It thus appears possible that a pathway that normally exerts inhibitory influences on water intake may course through ventromedial aspects of the brainstem.

It is interesting to note, in this connection, that the hyperphagia produced by chemical and electrolytic brainstem lesions that interrupt the ventral noradrenergic bundle (Ahlskog, Hoebel, & Breisch, 1974), as well as the feeding response to norepinephrine injections into the region of the

paraventricular nucleus (Leibowitz, Note 3), can be prevented by hypophysectomy. The numerous recent studies implicating the ventral noradrenergic bundle and its diencephalic projections in the regulation of food intake are particularly interesting in this context, because noradrenergic components of the hypothalamus have been implicated in the release of steroid hormones. For instance, steroids alter brain monoamine metabolism (Anton-Tay & Wurtman, 1971; Wurtman, 1971); hypothalamic norepinephrine levels vary with the estrous cycle (Stefano & Donoso, 1967); and norepinephrine levels, synthesis, and "turnover" have been shown to increase after ovariectomy (Anton-Tay, Anton, & Wurtman, 1973; Anton-Tay, Pelham, & Wurtman, 1969; Anton-Tay & Wurtman, 1968; Bapna, Neff, & Costa, 1971). More directly to the point under discussion, Simpson and DiCara (1973) have reported that noradrenergic stimulation of the anterior hypothalamus increased food intake regardless of plasma estradiol levels, whereas dopamine injections elicited feeding only when estradiol titers were low. From this and related observations, Simpson and DiCara (1973) derived the suggestion that high levels of estradiol might inhibit eating by decreasing hypothalamic dopamine-beta-hydroxylase, the enzyme necessary for the synthesis of norepinephrine from dopamine. Some experimental observations (e.g., Stern & Zwick, 1973) are not readily compatible with such an interpretation, but the possibility of a tie-in of the noradrenergic components of the hypothalamus, which have been implicated in feeding by numerous other lines of investigation discussed in this article, with the influence of steroid hormones is an intriguing possibility, which deserves further attention.

EVIDENCE FOR HYPOTHALAMIC CONTROL

Now that we have reviewed some of the evidence that questions the unique role of the hypothalamus in the regulation of ingestive behavior, it seems only fair to ask what evidence might be cited in support of the hypothesis that cellular components

in the region of the postulated feeding and satiety centers may, after all, contribute significantly to the regulation of food or water intake.

Electrical stimulation in the hypothalamus has been shown to elicit feeding or drinking (Andersson & Wyrwicka, 1957; Miller, 1957; Mogenson & Stevenson, 1967), as well as hunger- and thirst-motivated instrumental behaviors (Andersson & Wyrwicka, 1957; Coons, Levak, & Miller, 1965; Miller, 1963). The behavioral specificity of these elicited behaviors has been questioned (Cox et al., 1969; see Mogenson, 1973; Teitelbaum, 1973; Valenstein, 1973, for reviews of the current status of this issue), but even if one accepts the view, favored by most investigators in this field, that anatomically distinct hypothalamic mechanisms exist for the regulation of food and water intake, nothing indicates that the effects of electrical stimulation are due to an activation of cellular components of the region rather than fibers of passage.

Many years ago, I attempted to obtain evidence for feeding- and drinking-related neurons in the hypothalamus by observing the behavioral effects of microinjections of putative transmitters, which should act selectively on postsynaptic cellular membranes (Grossman, 1960, 1962a, 1962b). These experiments demonstrated that injections of norepinephrine and related compounds into the hypothalamus elicited feeding and inhibited drinking, whereas injections of acetylcholine and related synthetic cholinomimetics elicited drinking and inhibited feeding. These observations have been repeated and extended by others (e.g., Booth, 1967; Fisher & Coury, 1962; Leibowitz, 1970; Margules, 1970; Margules, Lewis, Dragovich, & Margules, 1972). It now appears that adrenergic compounds may not only elicit feeding but also have inhibitory effects on intake (Margules, 1970), depending, perhaps, on the injection site (Leibowitz, 1970, 1971), taste factors (Margules & Dragovich, 1973), or time of day (Margules et al., 1972). But is this persuasive evidence for the existence of neurons which are specific-

ally related to feeding, drinking, or satiety in the hypothalamus?

Drinking can be elicited by injections of cholinomimetics into numerous sites outside the hypothalamus (Fisher & Coury, 1962; Grossman, 1964a, 1964b), and the effect can be blocked by injections of anticholinergic agents, such as atropine, into many widely distributed sites in the limbic system (Levitt & Fisher, 1966). The ubiquity of the effective placements has suggested to some investigators that the compounds might diffuse to a single site of action, perhaps via the ventricles (Routtenberg, 1967). Autoradiographic data on the rate and extent of diffusion of atropine from an injection site (Grossman & Stumpf, 1969) are not readily congruent with such an interpretation, but the available data do not rule out the hypothesis, because *effective* diffusion and the effects of transport via the vascular system are not well understood at this time. The possibility that the high concentrations used in these experiments may produce nonspecific effects on fibers of passage has also not been ruled out.

My original investigation of the adrenergic feeding effect suggested that the most effective injections involved the perifornical region of the hypothalamus, but active sites were also seen rostral, dorsal, and medial to that area (Grossman, 1962a). Booth (1967) has reported a concentration of positive points in the dorsal regions of the anterior hypothalamus and basal thalamus and has suggested that projections from the stria medullaris might be responsible for the effect. Leibowitz (1970) noted that most positive injection sites were concentrated in the medial hypothalamus and has recently (Note 3) specifically implicated the paraventricular nucleus and the supraoptic nucleus. Feeding responses to alpha adrenergic compounds have not been reported from injection sites in the dorsolateral hypothalamus, where lesions result in persistent aphagia and adipsia. It may be significant, in this context, that the inhibitory effects that Margules and his associates (Margules, 1970; Margules & Dragovich, 1973; Mar-

gules et al., 1972) have observed were obtained from cannula placements in the lateral hypothalamus.

On balance, these observations are compatible with the notion that norepinephrine may act as an inhibitory transmitter on cellular components of the medial hypothalamus that exert inhibitory (i.e., satiety) influences on feeding. Such an interpretation is congruent with the lesion literature, but it is difficult to completely rule out the possibility of nonspecific drug effects that might involve fibers of passage rather than postsynaptic membranes, even though recent investigations have succeeded in eliciting (Leibowitz, Note 3) or facilitating (Ritter & Epstein, Note 2) feeding with nanogram doses of norepinephrine, which begin to approximate physiological quantities.

The strongest evidence that noradrenergic components of the hypothalamus play an important role in the regulation of food intake, at least in the monkey, has been reported by Yaksh and Myers (1972). Using a push-pull cannula, these investigators found that a perfusate of the hypothalamus of a food-deprived monkey would elicit feeding when injected into the hypothalamus of a sated animal and that a perfusate from the hypothalamus of a sated monkey would inhibit feeding in a deprived one. These observations are particularly significant in the context of this discussion, in view of the fact that perfusates from hungry monkeys contained an increased quantity of a norepinephrine-like substance and that injections of exogenous norepinephrine into the hypothalamus also elicited feeding. Unfortunately, these intriguing experiments have not yet been replicated in the rat, which has been the subject of most other experimental investigations in this field.

There are other experimental observations which support the hypothesis that cellular components of the hypothalamus regulate food and water intake but most admit to alternative explanations. Thus, the toxic compound goldthioglucose concentrates in the medial hypothalamus after systemic administration, and the resulting

lesion produces hyperphagia (Mayer & Marshall, 1956). The effect appears to be insulin dependent (Debons, Krimsky, From, & Cloutier, 1969) and is often interpreted to support the hypothesis that the ventromedial satiety center uses glucoreceptors to meter the nutritional state of the organism. This is certainly a plausible explanation, but a number of observations suggest that it is not compelling. For instance, goldthioglucose also concentrates (and produces lesions) in many other parts of the brain (see Liebelt & Perry, 1967, for review), suggesting that uptake may be determined by relative metabolic activity rather than specific glucoreceptor functions. Moreover, insulin, which is thought to elicit feeding by decreasing blood sugar, or more specifically, the utilization of sugar by the hypothesized glucoreceptors, continues to increase feeding in rats with extensive lesions in the ventromedial hypothalamus (Epstein & Teitelbaum, 1967). The importance of central glucoreceptors has been questioned by recent observations, which indicate that feeding responses to intracellular glucoprivation are abolished by lesions in the anterior hypothalamus (Blass & Kraly, 1974), zona incerta (Walsh & Grossman, unpublished observation, 1974), and midbrain tegmentum (Alheid & Grossman, unpublished observation, 1974) that do not modify ad-libitum food intake.

Additional evidence for the presence of feeding-related cells in the hypothalamus comes from investigations of the activity of single neurons in the region (see Oomura, 1973, for review). This technique can distinguish potentials generated by cells from those propagated by fibers of passage and thus provides evidence directly related to our question. Earlier investigations (Anand, Chhina, Sharma, Dua, & Singh, 1964; Desiraju, Banerjee, & Anand, 1968; Oomura, Kimura, Ooyama, Maeno, Iki, & Kuniyoshi, 1964) reported apparently systematic changes in single cell activity in the ventromedial and lateral hypothalamus following systemically induced hyperglycemia or hypoglycemia, cellular glucoprivation, or stimulation of gastric stretch receptors. These observations lend some

support to the hypothesis that glucoreceptors and receptors sensitive to distension of the stomach may be present in the hypothalamus but also admit other plausible interpretations, because the treatments were, in most cases, quite extreme (e.g., the hypothalamus has been implicated in "affective" reactions to noxious stimulation, and the observed changes in cellular activity may well have been related to such reactions).

More recently, insulin and glucose have been applied iontophoretically to individual cells in the ventromedial and lateral hypothalamus (e.g., Oomura, Ooyama, Yamamoto, Ono, & Kobayashi, 1969), and cells have been found in either location which increased or decreased their firing rate in response to such treatments. It is possible that the selective activation of a cell by glucose or insulin may not reflect reactions peculiar to feeding-related glucoreceptors, but these experiments do avoid some of the ambiguities inherent in the earlier studies.

Also in harmony with the hypothalamic model of food intake regulation is the observation that a reciprocal relationship appears to characterize at least some cells in the medial and lateral hypothalamus. Although the number of such concurrent observations is not large, and puzzling negative results have been published (Murphy & Renaud, 1968, 1969), it may be significant that an increase in the firing rate of a particular cell in the lateral hypothalamus is, at least under some circumstances, accompanied by a simultaneous shift to slower discharge rates of a cell in the ventromedial region (Oomura, Ooyama, Yamamoto, & Naka, 1967) and that electrical stimulation of one of these areas can inhibit the activity of a cell in the other (Van Atta & Sutin, 1971). What is missing, at this point, is a clear relationship between these phenomena and food deprivation, feeding, or satiety. Reciprocal changes in single cell activity have been seen mainly during various sleep stages and during the transition from sleep to wakefulness or from drowsiness to arousal, and their relationship to feeding is tenuous at best.

CONCLUSIONS

Where, then, do we stand in our understanding of the hypothalamic mechanisms believed to regulate food and water intake? It has been the thesis of the present review that the fiber systems that course through the hypothalamus may play a more important role in the etiology of the dysfunctions in food and water intake that are seen after hypothalamic lesions and stimulation than the widely accepted model of hypothalamic regulation implies. These considerations do not compel the interpretation that cellular components of the hypothalamus may not exercise important and possibly essential regulatory functions, but they do suggest that a re-examination of widely held and possibly unfounded preconceptions is in order.

Although the picture that has emerged from recent investigations is incomplete, puzzling, and in some instances not entirely consistent, it appears likely that the aphagia syndrome, which characterizes rats with lateral hypothalamic lesions, may be due entirely or in part to an interruption of interconnections between the striatum and lower brainstem, including the dopaminergic nigrostriatal projections, which have no known synapse in the lateral hypothalamus. The functional importance of the regulatory deficit (i.e., lack of response to cellular glucoprivation), which persists after recovery of voluntary feeding, has been questioned by the results of recent investigations. It therefore appears increasingly possible that the initial impairment in feeding may be the result of sensorimotor disturbances rather than a reflection of a disruption of motivational mechanisms specifically related to eating.

The effects of lateral hypothalamic lesions on water intake appear to be due to the destruction of tissues in the zona incerta. The resulting adipsia is not explained by sensory or motor dysfunctions, since feeding-related drinking occurs in animals that are entirely adipsic during periods of food deprivation and fail to respond to cellular dehydration or extracellular hypovolemia. There is some

evidence that the zona incerta and lateral preoptic region may contain receptors, which may "meter" the state of cellular and perhaps extracellular hydration. The role of fiber systems that course through the zona incerta and adjacent subthalamic region remains to be elucidated.

The results of many recent experiments indicate that the overeating and obesity seen after medial hypothalamic damage may be related at least in part to noradrenergic pathways, which may synapse on cells in the area.

REFERENCE NOTES

1. Simpson, J. B., & Epstein, A. N. *Subfornical organ lesions abolish drinking induced by systemic angiotensin-II*. Paper presented at the Satellite Symposium on the Physiology of Food and Fluid Intake, 26th International Congress of Physiological Sciences, Jerusalem, 1974.
2. Ritter, R. C., & Epstein, A. N. *Brain noradrenergic receptors may control meal size*. Paper presented at the Satellite Symposium on the Physiology of Food and Fluid Intake, 26th International Congress of Physiological Sciences, Jerusalem, 1974.
3. Leibowitz, S. F. *Catecholamines in the paraventricular and supraoptic nuclei: A possible function in regulating food ingestion and body water balance*. Paper presented at the Satellite Symposium on the Physiology of Food and Fluid Intake, 26th International Congress of Physiological Sciences, Jerusalem, 1974.
4. Ungerstedt, U. *Neuropharmacology of the control of food intake*. Paper presented at the Satellite Symposium on the Physiology of Food and Fluid Intake, 26th International Congress of Physiological Sciences, Jerusalem, 1974.

REFERENCES

- Ahlskog, J. E., & Hoebel, B. G. Overeating and obesity from damage to a noradrenergic system in the brain. *Science*, 1973, 182, 166-168.
- Ahlskog, J. E., Hoebel, B. G., & Breisch, S. T. Hyperphagia following lesions of the noradrenergic pathway is prevented by hypophysectomy. *Feeding Proceedings*, 1974, 33, 463. (Abstract)
- Albert, D. J., & Storlien, L. H. Hyperphagia in rats with cuts between the ventromedial and lateral hypothalamus. *Science*, 1969, 165, 599-600.
- Albert, D. J., Storlien, L. H., Wood, D., & Ehman, G. Further evidence for a complex system controlling feeding behavior. *Physiology & Behavior*, 1970, 5, 1075-1082.
- Anand, B. K., & Brobeck, J. R. Hypothalamic control of food intake in rats and cats. *Yale Journal of Biology and Medicine*, 1951, 24, 123-140.

- Anand, B. K., Chhina, G. S., Sharma, K. N., Dua, S., & Singh, B. Activity of single neurons in the hypothalamic feeding centers: Effect of glucose. *American Journal of Physiology*, 1964, 207, 1146-1154.
- Andén, N.-E., Fuxe, K., Hamberger, B., & Hökfelt, T. A quantitative study on the nigrostriatal dopamine neuron system in the rat. *Acta Physiologica Scandinavica*, 1966, 67, 306-312.
- Andersson, B. Invited comment: Osmoreceptors versus sodium receptors. In A. N. Epstein, H. R. Kissileff, & E. Stellar (Eds.), *The neuropsychology of thirst*. Washington, D. C.: V. H. Winston, 1973.
- Andersson, B., & Wyrwicka, W. The elicitation of a drinking motor conditioned reaction by electrical stimulation of the "drinking area" of the goat. *Acta Physiologica Scandinavica*, 1957, 41, 194-198.
- Anton-Tay, F., Anton, S. M., & Wurtman, R. J. Mechanism of changes in brain norepinephrine metabolism after ovariectomy. *Neuroendocrinology*, 1973, 6, 265-273.
- Anton-Tay, F., Pelham, R. W., & Wurtman, R. J. Increased turnover of ³H-norepinephrine in rat brain following castration or treatment with ovine follicle stimulating hormone. *Endocrinology*, 1969, 84, 1489-1492.
- Anton-Tay, F., & Wurtman, R. J. Norepinephrine: Turnover in rat brains after gonadectomy. *Science*, 1968, 159, 1245.
- Anton-Tay, F., & Wurtman, R. J. Brain monoamines and endocrine function. In L. Martini & W. Ganong (Eds.), *Frontiers in neuroendocrinology*. New York: Oxford, 1971.
- Arees, E. A., & Mayer, J. Anatomical connections between medial and lateral regions of the hypothalamus concerned with food intake. *Science*, 1967, 157, 1574-1575.
- Balagura, S., & Devenport, L. D. Feeding patterns of normal and ventromedial hypothalamic lesioned male and female rats. *Journal of Comparative and Physiological Psychology*, 1970, 71, 357-364.
- Balagura, S., Wilcox, R. H., & Coscina, D. V. The effect of diencephalic lesions on food intake and motor activity. *Physiology & Behavior*, 1969, 4, 629-633.
- Bapna, J., Neff, N. H., & Costa, E. A method for studying norepinephrine and serotonin metabolism in small regions of rat brain: Effect of ovariectomy on amine metabolism in anterior and posterior hypothalamus. *Endocrinology*, 1971, 89, 1345-1349.
- Berger, B. D., Wise, C. D., & Stein, L. Norepinephrine: Reversal of anorexia in rats with lateral hypothalamic damage. *Science*, 1971, 172, 281-284.
- Black, S. L., & Mogenson, G. J. The regulation of serum sodium in septal lesioned rats: A test of two hypotheses. *Physiology & Behavior*, 1973, 10, 379-384.
- Blass, E. M., & Epstein, A. N. A lateral preoptic osmosensitive zone for thirst in the rat. *Journal of Comparative and Physiological Psychology*, 1971, 76, 378-394.
- Blass, E. M., & Hanson, D. G. Primary hyperdipsia in the rat following septal lesions. *Journal of Comparative and Physiological Psychology*, 1970, 70, 87-93.
- Blass, E. M., & Kraly, F. S. Medial forebrain bundle lesions: Specific loss of feeding to decreased glucose utilization in rats. *Journal of Comparative and Physiological Psychology*, 1974, 86, 679-692.
- Blass, E. M., Nussbaum, A. I., & Hanson, D. G. Septal hyperdipsia: Specific enhancement of drinking to angiotensin in rats. *Journal of Comparative and Physiological Psychology*, 1974, 87, 422-439.
- Blatt, B., & Lyon, M. The interrelationship of forebrain and midbrain structures involved in feeding behavior. *Acta Neurologica Scandinavica*, 1968, 44, 576-595.
- Booth, D. A. Localization of the adrenergic feeding system in the rat diencephalon. *Science*, 1967, 158, 515-517.
- Brobeck, J. R. Mechanisms of the development of obesity in animals with hypothalamic lesions. *Physiological Reviews*, 1946, 26, 541-559.
- Brobeck, J. R., Wheatland, M., & Strominger, J. L. Variations in regulation of energy exchange associated with estrus, diestrus, and pseudopregnancy in rats. *Endocrinology*, 1947, 40, 65-72.
- Campbell, B. A., & Baez, L. A. Dissociation of arousal and regulatory behaviors following lesions of the lateral hypothalamus. *Journal of Comparative and Physiological Psychology*, 1974, 87, 142-149.
- Chase, P., & Moore, R. Y. Medial forebrain bundle and dorsomedial tegmentum lesions: Effects on operant behavior and activity in the rat. *Communications in Behavioral Biology* (Pt. A), 1968, 1, 133-141.
- Chi, C. C. Afferent connections of the ventromedial hypothalamus in the rat. *Brain Research*, 1970, 17, 439-445.
- Coons, E. E., Levak, M., & Miller, N. E. Lateral hypothalamus: Learning of food-seeking response motivated by electrical stimulation. *Science*, 1965, 150, 1320-1321.
- Corbit, J. D. Hyperphagic hyperreactivity to adulteration of drinking water with quinine HCl. *Journal of Comparative and Physiological Psychology*, 1965, 60, 123-124.
- Cosina, D. V., & Balagura, S. Avoidance and escape behavior of rats with aphagia produced by basal diencephalic lesions. *Physiology & Behavior*, 1970, 5, 651-658.
- Coury, J. N. Neural correlates of food and water intake. *Science*, 1967, 156, 1763-1765.
- Cox, V. C., & Kakolewski, J. W. Sex differences in body weight regulation in rats following hypothalamic lesions. *Communications in Behavioral Biology*, 1970, 5, 195-197.
- Cox, V. C., Kakolewski, J. W., & Valenstein, E. S. Ventromedial hypothalamic lesions and changes in body weight and food consumption in male and female rats. *Journal of Comparative and Physiological Psychology*, 1969, 67, 320-326.

- Dahlström, A., & Fuxe, K. Evidence for the existence of monoamine-containing neurons in the central nervous system. *Acta Physiologica Scandinavica* (Suppl. 232), 1964, 62, 1-55.
- Debons, A. F., Krinsky, I., From, A., & Cloutier, R. J. Rapid effects of insulin on the hypothalamic satiety center. *American Journal of Physiology*, 1969, 217, 1114-1118.
- Desiraju, T., Banerjee, M. G., & Anand, B. K. Activity of single neurons in the hypothalamic feeding centers: Effect of 2-deoxy-D-glucose. *Physiology & Behavior*, 1968, 3, 757-760.
- DiCara, L. V., & Wolf, G. Bar pressing for food reinforcement after lesions of efferent pathways from lateral hypothalamus. *Experimental Neurology*, 1968, 21, 231-235.
- Eager, R. P., Chi, C. C., & Wolf, G. Lateral hypothalamic projections to the hypothalamic ventromedial nucleus in the albino rat: Demonstration by means of simplified ammoniacal silver degeneration method. *Brain Research*, 1971, 29, 128-132.
- Epstein, A. N. The lateral hypothalamic syndrome: Its implications for the physiological psychology of hunger and thirst. In E. Stellar & J. M. Sprague (Eds.), *Progress in physiological psychology* (Vol. 4). New York: Academic Press, 1971.
- Epstein, A. N. Epilogue: Retrospect and prognosis. In A. N. Epstein, H. R. Kissileff, & E. Stellar (Eds.), *The neuropsychology of thirst*. Washington D. C.: V. H. Winston, 1973.
- Epstein, A. N., Fitzsimons, J. T., & Rolls, B. J. Drinking induced by injections of angiotensin into the brain of the rat. *Journal of Physiology*, 1970, 210, 457-474.
- Epstein, A. N., & Teitelbaum, P. Severe and persistent deficits in thirst produced by lateral hypothalamic damage. In M. J. Wayner (Ed.), *Thirst in the regulation of body water*. Oxford: Pergamon Press, 1964.
- Epstein, A. N., & Teitelbaum, P. Specific loss of the hypoglycemic control of feeding in recovered lateral rats. *American Journal of Physiology*, 1967, 213, 1159-1167.
- Fibiger, H. C., Zis, A. P., & McGeer, E. G. Feeding and drinking deficits after 6-hydroxydopamine administration in the rat: Similarities to the lateral hypothalamic syndrome. *Brain Research*, 1973, 55, 135-148.
- Fisher, A. E. Relationships between cholinergic and other dipsogens in the central mediation of thirst. In A. N. Epstein, H. R. Kissileff, & E. Stellar (Eds.), *The neuropsychology of thirst*. Washington, D.C.: V. H. Winston, 1973.
- Fisher, A. E., & Coury, J. N. Cholinergic tracing of a central neural circuit underlying the thirst drive. *Science*, 1962, 138, 691-693.
- Friedman, M. I., & Campbell, B. A. Ontogeny of thirst in the rat: Effects of hypertonic saline, polyethylene glycol, and vena cava ligation. *Journal of Comparative and Physiological Psychology*, 1974, 87, 37-46.
- Fuxe, K. Evidence for the existence of monoamine-containing neurons in the central nervous system: IV. Distribution of monoamine nerve terminals in the central nervous system. *Acta Physiologica Scandinavica* (Suppl. 247), 1965, 64, 37-85.
- Fuxe, K., & Hökfelt, T. Catecholamines in the hypothalamus. In F. Ganong & L. Martini (Eds.), *Frontiers in neuroendocrinology*. New York: Oxford University Press, 1969.
- Glick, S. D., Greenstein, S., & Waters, D. H. Lateral hypothalamic lesions and striatal dopamine levels. *Life Sciences*, 1974, 14, 747-750.
- Gold, R. M. Aphagia and adipisia following unilateral and bilaterally asymmetrical lesions in rats. *Physiology & Behavior*, 1967, 2, 211-220.
- Gold, R. M. Hypothalamic hyperphagia: Males get just as fat as females. *Journal of Comparative and Physiological Psychology*, 1970, 71, 347-356.
- Gold, R. M. Hypothalamic obesity: The myth of the ventromedial nucleus. *Science*, 1973, 182, 488-490.
- Gold, R. M., Quackenbush, P. M., & Kapatos, G. Obesity following combination of rostralateral to VMH cut and contralateral mammillary area lesions. *Journal of Comparative and Physiological Psychology*, 1972, 79, 210-218.
- Graff, H., & Stellar, E. Hyperphagia, obesity and finickiness. *Journal of Comparative and Physiological Psychology*, 1962, 55, 418-424.
- Grossman, S. P. Eating or drinking elicited by direct adrenergic or cholinergic stimulation of hypothalamus. *Science*, 1960, 132, 301-302.
- Grossman, S. P. Direct adrenergic and cholinergic stimulation of hypothalamic mechanisms. *American Journal of Physiology*, 1962, 202, 872-882. (a)
- Grossman, S. P. Effects of adrenergic and cholinergic blocking agents on hypothalamic mechanisms. *American Journal of Physiology*, 1962, 202, 1230-1236. (b)
- Grossman, S. P. Behavioral effects of chemical stimulation of the ventral amygdala. *Journal of Comparative and Physiological Psychology*, 1964, 57, 29-36. (a)
- Grossman, S. P. Effects of chemical stimulation of the septal area on motivation. *Journal of Comparative and Physiological Psychology*, 1964, 58, 194-200. (b)
- Grossman, S. P. Changes in food and water intake associated with an interruption of the anterior or posterior fiber connections of the hypothalamus. *Journal of Comparative and Physiological Psychology*, 1971, 75, 23-31.
- Grossman, S. P. Neurophysiologic aspects: Extrahypothalamic factors in the regulation of food intake. *Advances in Psychosomatic Medicine*, 1972, 7, 49-72.
- Grossman, S. P., & Grossman, L. Food and water intake following lesions or electrical stimulation of the amygdala. *American Journal of Physiology*, 1963, 205, 761-765.
- Grossman, S. P., & Grossman, L. Surgical interruption of the anterior or posterior connections of the hypothalamus: Effects on aggressive and avoidance behavior. *Physiology & Behavior*, 1970, 5, 1313-1317.

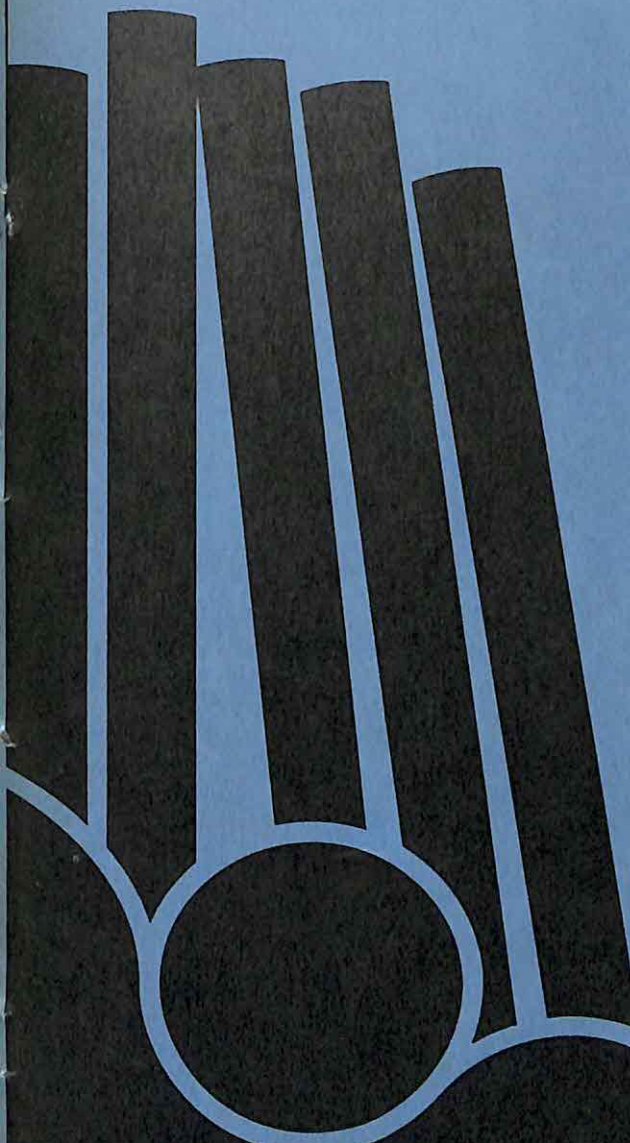
- Grossman, S. P., & Grossman, L. Food and water intake in rats with parasagittal knife cuts medial or lateral to the lateral hypothalamus. *Journal of Comparative and Physiological Psychology*, 1971, 74, 148-156.
- Grossman, S. P., & Grossman, L. Persisting deficits in rats "recovered" from transections of fibers which enter or leave hypothalamus laterally. *Journal of Comparative and Physiological Psychology*, 1973, 85, 515-527.
- Grossman, S. P., Grossman, L., & Walsh, L. L. Functional organization of the rat amygdala with respect to avoidance behavior. *Journal of Comparative and Physiological Psychology*, 1975, 88, 829-850.
- Grossman, S. P., & Stumpf, W. E. Intracranial drug implants: An autoradiographic analysis of diffusion. *Science*, 1969, 166, 1410-1412.
- Guillery, R. W. Degeneration in the hypothalamic connexions of the albino rat. *Journal of Anatomy*, 1957, 91, 91-115.
- Harrell, L. E., & Balagura, S. The influence of ovarian hormones on the recovery period following lateral hypothalamic lesions. *Journal of Comparative and Physiological Psychology*, 1975, 88, 194-201.
- Harvey, J. A., & Hunt, H. F. Effect of septal lesions on thirst in the rat as indicated by water consumption and operant responding for water reward. *Journal of Comparative and Physiological Psychology*, 1965, 59, 49-56.
- Hervey, G. R., & Hutchinson, I. The effects of testosterone on body weight and composition in the rat. *Journal of Endocrinology*, 1973, 57, 24.
- Hetherington, A. W., & Ranson, S. W. Hypothalamic lesions and adiposity in the rat. *Anatomical Record*, 1940, 78, 149-172.
- Hetherington, A. W., & Ranson, S. W. The relation of various hypothalamic lesions to adiposity in the rat. *Journal of Comparative Neurology*, 1942, 76, 475-499.
- Hoebel, B. G., & Teitelbaum, P. Weight regulation by normal and hypothalamic hyperphagic rats. *Journal of Comparative and Physiological Psychology*, 1966, 61, 189-193.
- Jankowiak, R., & Stern, J. J. Food intake and body weight modifications following medial hypothalamic hormone implants in female rats. *Physiology & Behavior*, 1974, 12, 875-879.
- Johnson, A. K. The role of the cerebral ventricular system in angiotensin-induced thirst. In J. T. Fitzsimons, G. Peters, & L. Peters-Haefeli (Eds.), *Control mechanisms of drinking*. New York: Springer, in press.
- Joseph, S. A., & Knigge, J. M. Effects of VMH lesions in adult and newborn guinea pigs. *Neuroendocrinology*, 1968, 3, 309-331.
- Kaelber, W. W., & Leeson, C. R. A degeneration and electron microscope study of the nucleus hypothalamicus ventromedialis of the rat. *Journal of Anatomy*, 1967, 101, 209-221.
- Kakolewski, J. W., Cox, V. C., & Valenstein, E. S. Sex differences in body-weight changes following gonadectomy of rats. *Psychological Reports*, 1968, 22, 547-554.
- Kennedy, G. C. The development with age of hypothalamic restraint upon the appetite of the rat. *Journal of Endocrinology*, 1957, 16, 9-17.
- Kennedy, G. C., & Mitra, J. Hypothalamic control of energy balance and the reproductive cycle in the rat. *Journal of Physiology*, 1963, 166, 395-407.
- Kenney, N. J., & Mook, D. G. Effects of ovariectomy on meal pattern in the albino rat. *Journal of Comparative and Physiological Psychology*, 1974, 87, 302-309.
- Kent, E. W., & Grossman, S. P. Elimination of learned behaviors after transection of fibers crossing the lateral border of the hypothalamus. *Physiology & Behavior*, 1973, 10, 953-963.
- Kent, E. W., Rezak, M., & Grossman, S. P. Transection and chemical lesion of nigrostriatal pathways: A comparison of effects on learned behavior. *Proceedings of the Society for Neuroscience*, 1973, 3, 410. (Abstract)
- King, J. M., & Cox, V. C. The effects of estrogens on food intake and body weight following ventromedial hypothalamic lesions. *Physiological Psychology*, 1973, 1, 261-264.
- Kissileff, H. R., & Epstein, A. N. Exaggerated prandial drinking in the "recovered lateral" rat without saliva. *Journal of Comparative and Physiological Psychology*, 1969, 67, 301-308.
- Knook, H. L. *The fibre connections of the forebrain*. Amsterdam: Royal Van Gorcum, 1965.
- Lammers, H. J. The neural connections of the amygdaloid complex in mammals. In B. E. Eleftheriou (Ed.), *The neurobiology of the amygdala*. New York: Plenum Press, 1972.
- Laursen, A. Corpus striatum. *Acta Physiologica Scandinavica* (Suppl. 211), 1963, 59, 1-106.
- Lehr, D. Invited comment: Comments to papers on "thirst" by Drs. Fisher, Harvey, and Setler. In A. N. Epstein, H. R. Kissileff, & E. Stellar (Eds.), *The neurophysiology of thirst*. Washington, D.C.: V. H. Winston, 1973.
- Leibowitz, S. F. Reciprocal hunger-regulating circuits involving alpha- and beta-adrenergic receptors located, respectively, in the ventromedial and lateral hypothalamus. *Proceedings of the National Academy of Sciences of the United States of America*, 1970, 67, 1063-1070.
- Leibowitz, S. F. Hypothalamic alpha- and beta-adrenergic systems regulate both thirst and hunger in the rat. *Proceedings of the National Academy of Sciences of the United States of America*, 1971, 68, 332-334.
- Levine, M. S., Ferguson, N., Kreinick, C. J., Gustafson, J. W., & Schwartzbaum, J. S. Sensorimotor dysfunctions and aphagia and adipsia following pallidal lesions in rats. *Journal of Comparative and Physiological Psychology*, 1971, 77, 282-293.
- Levine, M. S., & Schwartzbaum, J. Sensorimotor functions of the striatopallidal system and lateral hypothalamus and consummatory behavior in

- rats. *Journal of Comparative and Physiological Psychology*, 1973, 85, 615-635.
- Levitt, R. A., & Fisher, A. E. Anticholinergic blockade of centrally induced thirst. *Science*, 1966, 154, 520-522.
- Liebelt, R. A., & Perry, J. H. Action of goldthioglucose on the central nervous system. In C. F. Code (Ed.), *Handbook of physiology. Section 6: Alimentary canal* (Vol. 1). Washington, D.C.: American Physiological Society, 1967.
- Lubar, J. F., Boyce, B. A., & Schaefer, C. F. Etiology of polydipsia and polyuria in rats with septal lesions. *Physiology & Behavior*, 1968, 3, 289-292.
- Lyon, M., Halpern, M., & Mintz, E. The significance of the mesencephalon for coordinated feeding behavior. *Acta Neurologica Scandinavica*, 1968, 44, 323-346.
- Mackay, E. M., Calloway, J. W., & Barnes, R. H. Hyperalimination in normal animals produced by protamine insulin. *Journal of Nutrition*, 1940, 20, 59-66.
- Margules, D. L. Alpha-adrenergic receptors in hypothalamus for the suppression of feeding behavior by satiety. *Journal of Comparative and Physiological Psychology*, 1970, 73, 1-12.
- Margules, D. L., & Dragovich, J. Studies on phentolamine-induced overeating and finickiness. *Journal of Comparative and Physiological Psychology*, 1973, 84, 644-651.
- Margules, D. L., Lewis, M. J., Dragovich, J. A., & Margules, A. S. Hypothalamic norepinephrine: Circadian rhythms and the control of feeding behavior. *Science*, 1972, 178, 640-643.
- Marshall, J. F., Richardson, J. S., & Teitelbaum, P. Nigrostriatal bundle damage and the lateral hypothalamic syndrome. *Journal of Comparative and Physiological Psychology*, 1974, 87, 808-830.
- Marshall, J. F., & Teitelbaum, P. A comparison of the eating in response to hypothermic and glucoprivic challenges after nigral 6-hydroxydopamine and lateral hypothalamic lesions in rats. *Brain Research*, 1973, 55, 229-233.
- Marshall, J. F., & Teitelbaum, P. Further analysis of sensory inattention following lateral hypothalamic damage in rats. *Journal of Comparative and Physiological Psychology*, 1974, 86, 375-395.
- Marshall, J. F., Turner, B. H., & Teitelbaum, P. Sensory neglect produced by lateral hypothalamic damage. *Science*, 1971, 174, 523-525.
- Mayer, J., & Marshall, N. B. Specificity of goldthioglucose for ventromedial hypothalamic lesions and obesity. *Nature*, 1956, 178, 1399-1400.
- Miller, N. E. Experiments on motivation. *Science*, 1957, 126, 1271-1278.
- Miller, N. E. Some motivational effects of electrical and chemical stimulation of the brain. *Electroencephalography and Clinical Neurophysiology* (Suppl. 24), 1963, 15, 247-259.
- Millhouse, O. E. A Golgi study of the descending medial forebrain bundle. *Brain Research*, 1969, 15, 341-363.
- Millhouse, O. E. Certain ventromedial hypothalamic afferents. *Brain Research*, 1973, 55, 89-105. (a)
- Millhouse, O. E. The organization of the ventromedial hypothalamic nucleus. *Brain Research*, 1973, 55, 71-87. (b)
- Mogenson, G. J. Hypothalamic limbic mechanisms in the control of water intake. In A. N. Epstein, H. R. Kissileff, & E. Stellar (Eds.), *The neurophysiology of thirst*. Washington, D.C.: V. H. Winston, 1973.
- Mogenson, G. J., & Stevenson, J. A. F. Drinking induced by electrical stimulation of the lateral hypothalamus. *Experimental Neurology*, 1967, 17, 119-127.
- Montemurro, D. G. Inhibition of hypothalamic obesity in the mouse with diethylstilbestrol. *Canadian Journal of Physiology and Pharmacology*, 1971, 49, 554-558.
- Mook, D. G., Kenney, N. J., Roberts, S., Nussbaum, A. I., & Rodier, W. L., III. Ovarian-adrenal interactions in regulation of body weight by female rats. *Journal of Comparative and Physiological Psychology*, 1972, 81, 198-211.
- Morgane, P. J. Alterations in feeding and drinking behavior of rats with lesions in globi pallidi. *American Journal of Physiology*, 1961, 201, 420-428. (a)
- Morgane, P. J. Distinct "feeding" and "hunger motivating" systems in the lateral hypothalamus of the rat. *Science*, 1961, 133, 887-888. (b)
- Morgane, P. J. Medial forebrain bundle and "feeding centers" of the hypothalamus. *Journal of Comparative Neurology*, 1961, 117, 1-25. (c)
- Murphy, J. T., & Renaud, L. P. Inhibitory interneurons in the ventromedial nucleus of the hypothalamus. *Brain Research*, 1968, 9, 385-389.
- Murphy, J. T., & Renaud, L. P. Mechanisms of inhibition in the ventromedial nucleus of the hypothalamus. *Journal of Neurophysiology*, 1969, 32, 85-102.
- Nauta, W. J. H. Silver impregnation of degenerating axons. In W. F. Windle (Ed.), *New research techniques of neuroanatomy*. Springfield, Ill.: Charles C Thomas, 1957.
- Nauta, W. J. H. Hippocampal projections and related neural pathways to the midbrain in the cat. *Brain*, 1958, 81, 319-340.
- Oltmans, G. A., & Harvey, J. A. LH syndrome and brain catecholamine levels after lesions of the nigrostriatal bundle. *Physiology & Behavior*, 1972, 8, 69-78.
- Oomura, Y. Central mechanisms of feeding. *Advances in Biophysics*, 1973, 5, 65-136.
- Oomura, Y., Kimura, K., Ooyama, H., Maeno, T., Iki, M., & Kuniyoshi, M. Reciprocal activities of the ventromedial and lateral hypothalamic areas of cats. *Science*, 1964, 143, 484-485.
- Oomura, Y., Ooyama, H., Yamamoto, T., & Naka, F. Reciprocal relationship of the lateral and ventromedial hypothalamus in the regulation of food intake. *Physiology & Behavior*, 1967, 2, 97-115.

- Oomura, Y., Ooyama, H., Yamamoto, T., Ono, T., & Kobayashi, N. Behavior of hypothalamic unit activity during electrophoretic applications of drugs. *Annals of the New York Academy of Science*, 1969, 157, 642-665.
- Palka, Y., Coyer, D., & Critchlow, V. Effects of isolation of medial basal hypothalamus on pituitary-adrenal and pituitary-ovarian functions. *Neuroendocrinology*, 1969, 5, 333-349.
- Palka, Y., Liebelt, R. A., & Critchlow, V. Obesity and increased growth following partial or complete isolation of ventromedial hypothalamus. *Physiology & Behavior*, 1971, 7, 187-194.
- Parker, S. W., & Feldman, S. M. Effects of mesencephalic lesions on feeding behavior in rats. *Experimental Neurology*, 1967, 17, 313-326.
- Paxinos, G. The hypothalamus: Neural systems involved in feeding, irritability, aggression, and copulation in male rats. *Journal of Comparative and Physiological Psychology*, 1974, 87, 110-119.
- Paxinos, G., & Bindra, D. Hypothalamic knife cuts: Effects on eating, drinking, irritability, aggression, and copulation in the male rat. *Journal of Comparative and Physiological Psychology*, 1972, 79, 219-229.
- Paxinos, G., & Bindra, D. Hypothalamic and mid-brain neural pathways involved in eating, drinking, irritability, aggression, and copulation in rats. *Journal of Comparative and Physiological Psychology*, 1973, 82, 1-14.
- Peck, J. W., & Novin, D. Evidence that osmoreceptors mediating drinking in rabbits are in the lateral preoptic area. *Journal of Comparative and Physiological Psychology*, 1971, 74, 134-147.
- Pfaff, D. W. Sex differences in food intake changes following pituitary growth hormone or prolactin injections. *Proceedings of the 77th Annual Convention of the American Psychological Association*, 1969, 4, 211-212.
- Powley, T. L., & Keesey, R. E. Relationship of body weight to the lateral hypothalamic feeding syndrome. *Journal of Comparative and Physiological Psychology*, 1970, 70, 25-36.
- Powley, T. L., & Opsahl, C. A. Ventromedial hypothalamic obesity abolished by subdiaphragmatic vagotomy. *American Journal of Physiology*, 1974, 226, 25-33.
- Pubols, L. M. Changes in food-motivated behavior of rats as a function of septal and amygdaloid lesions. *Experimental Neurology*, 1966, 15, 240-254.
- Raisman, G. Neural connections of the hypothalamus. *British Medical Bulletin*, 1966, 22, 197-201.
- Redick, J. H., Nussbaum A. I., & Mook, D. Estradiol independence on body weight. *Physiology & Behavior*, 1973, 10, 543-557.
- Rehovsky, D. A., & Wampler, R. S. Failure to obtain sex differences in development of obesity following VMH lesions in rats. *Journal of Comparative and Physiological Psychology*, 1972, 78, 102-112.
- Richardson, J. S., & Jacobowitz, D. M. The effects of 6-hydroxydopa on brain catecholamines and behavior. *Federation Proceedings*, 1972, 31, 529.
- Richter, C. P., & Hawkes, C. D. Increased spontaneous activity and food intake produced in rats by removal of the frontal poles of the brain. *Journal of Neurology and Psychiatry*, 1939, 2, 231.
- Rodgers, W. L., Epstein, A. N., & Teitelbaum, P. Lateral hypothalamic aphagia: Motor failure or motivational deficit? *American Journal of Physiology*, 1965, 208, 334-342.
- Routtenberg, A. Drinking induced by carbachol: Thirst circuit or ventricular modification? *Science*, 1967, 157, 838-839.
- Runnels, P., & Thompson, R. Hypothalamic structures critical for the performance of a locomotor escape response in the rat. *Brain Research*, 1969, 13, 328-337.
- Schreiner, L., Rioch, McK., Pechtel, C., & Masserman, J. H. Behavioral changes following thalamic injury in cat. *Journal of Neurophysiology*, 1953, 16, 234-246.
- Scalfani, A. Neural pathways involved in the ventromedial hypothalamic lesion syndrome in the rat. *Journal of Comparative and Physiological Psychology*, 1971, 77, 70-96.
- Scalfani, A., & Grossman, S. P. Hyperphagia produced by knife cuts between the medial and lateral hypothalamus in the rat. *Physiology & Behavior*, 1969, 4, 533-538.
- Simpson, C. W., & DiCara, L. V. Estradiol inhibition of catecholamine elicited eating in the female rat. *Pharmacology Biochemistry and Behavior*, 1973, 1, 413-419.
- Simpson, J. B., & Routtenberg, A. Subfornical organ: Site of drinking elicitation by Angiotensin II. *Science*, 1973, 181, 1172-1175.
- Skultety, F. M., & Gary, T. M. Experimental hyperphagia in cats following destructive mid-brain lesions. *Neurology*, 1962, 12, 394-401.
- Smith, G. P., & Epstein, A. N. Increased feeding in response to decreased glucose utilization in the rat and monkey. *American Journal of Physiology*, 1969, 217, 1083-1087.
- Smith, G. P., Strohmayr, A. J., & Reis, D. J. Effect of lateral hypothalamic injections of 6-hydroxydopamine on food and water intake in rats. *Nature New Biology*, 1972, 235, 27-29.
- Stefano, F. J. E., & Donoso, A. O. Norepinephrine levels in the rat hypothalamus during the estrous cycle. *Endocrinology*, 1967, 81, 1405-1406.
- Stellar, E. The physiology of motivation. *Psychological Review*, 1954, 61, 5-22.
- Stern, J. J., & Zwick, G. Effects of intraventricular norepinephrine and estradiol benzoate on weight regulatory behavior in female rats. *Behavioral Biology*, 1973, 9, 605-612.
- Storlien, L. H., & Albert, D. J. The effect of VMH lesions, lateral cuts and anterior cuts on food intake, activity level, food motivation, and reactivity to taste. *Physiology & Behavior*, 1972, 9, 191-197.

- Stricker, E. M. Some physiological and motivational properties of the hypovolemic stimulus for thirst. *Physiology & Behavior*, 1968, 3, 379-385.
- Stricker, E. M. Thirst, sodium appetite, and complementary physiological contributions to the regulation of intravascular fluid volume. In A. N. Epstein, H. R. Kissileff, & E. Stellar (Eds.), *The neuropsychology of thirst*. Washington, D.C.: V. H. Winston, 1973.
- Stricker, E. M., & Wolf, G. The effect of hypovolemia on drinking in rats with lateral hypothalamic damage. *Proceedings of the Society for Experimental Biology and Medicine*, 1967, 124, 816-820.
- Stricker, E. M., & Zigmond, M. J. Effects on homeostasis of intraventricular injections of 6-hydroxydopamine in rats. *Journal of Comparative and Physiological Psychology*, 1974, 86, 973-994.
- Sutin, J., & Eager, R. P. Fiber degeneration following lesions in the hypothalamic ventromedial nucleus. *Annals of the New York Academy of Science*, 1969, 157, 610-628.
- Teitelbaum, P. Sensory control of hypothalamic hyperphagia. *Journal of Comparative and Physiological Psychology*, 1955, 48, 156-163.
- Teitelbaum, P. Motivation and control of food intake. In C. F. Code (Ed.), *Handbook of Physiology. Section 6: Alimentary canal* (Vol. 1). Washington, D.C.: American Physiological Society, 1967.
- Teitelbaum, P. Discussion: On the use of electrical stimulation to study hypothalamic structure and function. In A. N. Epstein, H. R. Kissileff, & E. Stellar (Eds.), *The neuropsychology of thirst*. Washington, D.C.: V. H. Winston, 1973.
- Teitelbaum, P., Cheng, M.-F., & Rozin, P. Development of feeding parallels its recovery after hypothalamic damage. *Journal of Comparative and Physiological Psychology*, 1969, 67, 430-441.
- Teitelbaum, P., & Epstein, A. N. The lateral hypothalamic syndrome: Recovery of feeding and drinking after lateral hypothalamic lesions. *Psychological Review*, 1962, 69, 74-90.
- Ungerstedt, U. Adipsia and aphagia after 6-hydroxydopamine induced degeneration of the nigrostriatal dopamine system. *Acta Physiologica Scandinavica* (Suppl. 367), 1971, 81, 95-122.
- Valenstein, E. S. Invited comment: Electrical stimulation and hypothalamic function: Historical perspective. In A. Epstein, H. R. Kissileff, & E. Stellar (Eds.), *The neuropsychology of thirst*. Washington, D.C.: V. H. Winston, 1973.
- Valenstein, E. S., Cox, V. C., & Kakolewski, J. W. Sex differences in hyperphagia and body weight following hypothalamic damage. *Annals of the New York Academy of Science*, 1969, 157, 1030-1048.
- Van Atta, L., & Sutin, J. The response of single lateral hypothalamic neurons to ventromedial nucleus and limbic stimulation. *Physiology & Behavior*, 1971, 6, 523-536.
- Wade, G. N. Gonadal hormones and behavioral regulation of body weight. *Physiology & Behavior*, 1972, 8, 523-534.
- Wade, G. N., & Zucker, I. Modulation of food intake and locomotor activity in female rats by diencephalic hormone implants. *Journal of Comparative and Physiological Psychology*, 1970, 72, 328-336.
- Walsh, L. L., & Grossman, S. P. Zona incerta lesions: Disruption of regulatory water intake. *Physiology & Behavior*, 1973, 11, 885-887.
- Walsh, L. L., & Grossman, S. P. Effects of zona incerta lesions and knife cuts on water intake following cellular and extracellular dehydration. *Proceedings of the Society for Neuroscience*, 1974, 4, 464. (Abstract)
- Wampler, R. S., Gier, H. T. Continuing gonadal function in rats with ventromedial hypothalamic area lesions. *Journal of Comparative and Physiological Psychology*, 1974, 87, 831-841.
- Wayner, M. J., Cott, A., Millner, J., & Tartaglione, R. Loss of 2-deoxy-D-glucose induced eating in recovered lateral rats. *Physiology & Behavior*, 1971, 7, 881-884.
- Webster, K. E. Cortico-striate interrelations in the albino rat. *Journal of Anatomy*, 1961, 95, 532-543.
- Webster, K. E. The cortico-striate projection in the cat. *Journal of Anatomy*, 1965, 99, 329.
- Wolf, G., & Sutin, J. Fiber degeneration after lateral hypothalamic lesions in the rat. *Journal of Comparative Neurology*, 1966, 127, 137-156.
- Wurtman, R. J. (Ed.). Brain monoamines and endocrine function. *Neurosciences Research Program Bulletin*, 1971, 9, 172-297.
- Yaksh, T. L., & Myers, R. D. Neurohumoral substance released from hypothalamus of the monkey during hunger and satiety. *American Journal of Physiology*, 1972, 222, 503-515.
- Zeigler, H. P., & Karten, H. S. Central trigeminal structures and the lateral hypothalamus syndrome in the rat. *Science*, 1974, 186, 636-637.
- Zigmond, M. J., & Stricker, E. M. Deficits in feeding behavior after intraventricular injections of 6-hydroxydopamine in rats. *Science*, 1972, 177, 1211-1213.
- Zucker, I. Hormonal determinants of sex differences in saccharin preference, food intake, and body weight. *Physiology & Behavior*, 1969, 4, 595-602.
- Zucker, I. Body weight and age as factors determining estrogen responsiveness in the rat feeding system. *Behavioral Biology*, 1972, 7, 527-542.

(Received September 16, 1974)



Publication Manual of the American Psychological Association

Second edition, 1974. Current guide to the preparation of manuscripts for APA and related journals. This edition adds many examples, including an entire sample paper. New sections on student papers, metrication, and mathematical copy. Enlarged sections on content and organization of an article, typing the manuscript, and all aspects of APA style, including reference lists and tables. 136 pp. \$3.

Please note that payment must accompany all orders amounting to \$15 or less except for those orders submitted on institutional purchase order forms.

American Psychological Association
Publication Sales Dept.
1200 17th Street, N.W.
Washington, D.C. 20036



2718
75

Thesaurus of Psychological Index Terms

Contains over 4,000 terms representing an authoritative, unambiguous, and hierarchically structured vocabulary used by psychology and related disciplines. This vocabulary will facilitate precise content representation of concepts described in psychological literature and allow for more flexibility in automated search and retrieval operations. These two factors promote more effective communication of psychological concepts in original works and research, ultimately aiding the content of publications to be focused in accordance with the interest patterns of psychologists.

This **Thesaurus**, developed by the Psychological Abstracts Information Services, is divided into three sections:

The **relationship section** shows the relationships among concepts that are associated with each other. Term relationships noted are: synonymous, broader, narrower, or related.

The **alphabetical section** presents each preferred term in alphabetical order, enabling rapid identification of terms used for indexing.

The **hierarchical section** contains 17 broad content areas in which terms are presented in descending order according to breadth of concept represented. The 17 categories correspond to major classifications used in *Psychological Abstracts*.

The relationship, alphabetical, and hierarchical sections will assist the user in selecting appropriate terms for different uses.

Administrators

- determine areas of research for an organization
- assure representative expertise in personnel selection

Researchers and Authors

- find synonyms for a term representing a concept
- select terms that represent a concept
- define subject matter topics for a paper or book

Educators

- structure a curriculum
- use as a tutorial aid

Consultants

- search the literature
- aid in retrieval of information

Librarians

- provide access to psychological literature
- aid in cataloging psychology material
- aid in ordering new psychological literature

Students

- locate concepts relevant to psychology
- focus areas of study

The **Thesaurus** contains 350 pages with a hard-cover and costs \$12. Please note that all orders amounting to \$15 or less must be accompanied by full payment except for those orders submitted on institutional purchase order forms.

American Psychological Association
Publication Sales Department
1200 17th Street, N.W.
Washington, D.C. 20036



#Kc/Ks
to the
Docket 2
Abraham
2/17/75

PSYCHOLOGICAL REVIEW

GEORGE MANDLER, EDITOR

JEAN M. MANDLER, ASSOCIATE EDITOR

UNIVERSITY OF CALIFORNIA, SAN DIEGO

Vol. 82, No. 4

July 1975

Articles

A Schema Theory of Discrete Motor Skill Learning
Richard A. Schmidt 225

Structure of Memory Traces
Endel Tulving and Michael J. Watkins 261

A Theory of Attention: Variations in the Associability of
Stimuli with Reinforcement
N. J. Mackintosh 276

An Emotion-Attribution Approach to Moral Behavior:
Interfacing Cognitive and Avoidance Theories of Moral
Development
Richard A. Dienstbier, Donald Hillman, John Lehnhoff,
Judith Hillman, and Maureen C. Valkenaar 299

Theoretical Note

A Testable Property of Some Discrete-State Models for
Recognition Memory
Dirk Vorberg and Rainer Schmidt 316

Editorial Consultants

The editor gratefully acknowledges the efforts of the following consultants who helped review manuscripts for this issue. They are part of a large group of psychologists who have kindly agreed to assist in the editing of manuscripts for *Psychological Review*.

WESLEY ALLINSMITH
JOHN R. ANDERSON
L. BERKOWITZ
INA M. BILODEAU
ROBERT A. BJORK
JOSEPH CAMPIONE
HERBERT H. CLARK
GEORGE COLLIER

RICHARD JAGACINSKI
H. M. JENKINS
EDWARD E. JONES
HAROLD H. KELLEY
MARK R. LEPPER
GERALD S. LEVENTHAL
GEORGE LEVINGER
RALPH MILLER
ANNE D. PICK

LARRY R. SQUIRE
JOHN E. R. STADDON
DANIEL STOKOLS
JOHN E. TAPLIN
WAYNE A. WICKELGREN
MORTON WIENER
AL YONAS
DAVID ZEAMAN

The *Psychological Review* publishes articles that make theoretical contributions to any area of scientific psychology. Preference is given to papers that advance theory rather than review it, and to statements that are specifically theoretical rather than programmatic. Surveys of the literature, problems of method and design, and papers that primarily report empirical findings are ordinarily not appropriate. Letters not exceeding 250 words on topics of theoretical interest should be submitted in duplicate.

Preparation of Manuscripts: Instructions for the preparation of manuscripts appear in the *Publication Manual of the American Psychological Association* (2nd ed.). Submit three copies; one copy must be the original typed copy. All parts of the manuscript must be double spaced. In order to facilitate blind review, the author's name should not appear on the manuscript or abstract, but each copy of the manuscript that identify the author should be typed on separate pages. Figures should be submitted as glossy prints ready for reproduction; duplicate figures may be photocopies. Metric equivalents must be given for all nonmetric measurements. Authors should keep a copy of the manuscript to guard against loss.

Abstracts: Manuscripts must be accompanied by an abstract of 100 to 175 words typed on a separate sheet of paper. Detailed instructions for preparation of abstracts appear in the *Publication Manual* or may be obtained from the Editor.

Manuscripts: Submit manuscripts in triplicate to the Editor, GEORGE MANDLER, Department of Psychology, University of California, San Diego, La Jolla, California 92037 according to instructions provided above.

Copyright and Permissions: All rights reserved. Written permission must be obtained from the American Psychological Association for copying or reprinting text of more than 500 words, tables, or figures. Permissions are normally granted contingent upon fee of \$10 per page of APA material. Permission fees are waived for authors who wish to reproduce their own articles. Address requests for permission to the Permissions Office, American Psychological Association, 1200 Seventeenth Street, N.W., Washington, D. C. 20036.

Subscriptions: Subscriptions are available on a calendar year basis only (January through December). Nonmember rates for 1975: \$15 domestic, \$16 foreign, \$3 single issue. APA member rate: \$6. Write to Subscription Section, APA.

Reprints: Authors may order reprints of their articles from the printer when they receive galley proofs.

Back Issues and Back Volumes: For information regarding back issues or back volumes write to Order Department, APA, Box 1017, Greenwich, Connecticut 06830; University Microfilms, Ann Arbor, Michigan 48106; or Princeton Microfilms, Princeton, New Jersey 08540.

Microform Editions: For information regarding microform editions write to any of the following: Johnson Associates, Inc., P.O. Box 1017, Greenwich, Connecticut 06830; University Microfilms, Ann Arbor, Michigan 48106; or Princeton Microfilms, Princeton, New Jersey 08540.

Change of Address: Send change of address notice and a recent mailing label to the attention of the Subscription Section, APA, 30 days prior to the actual change of address. Undelivered copies resulting from address changes will not be replaced; second-class mail is not forwarded unless subscribers notify the local post office in writing that they will guarantee second-class forwarding postage.

Published bimonthly (beginning in January) in one volume per year, six issues, by the American Psychological Association, Inc., 1200 Seventeenth Street, N.W., Washington, D. C. 20036. Printed in the U.S.A. Second-class postage paid at Washington, D. C., and additional mailing offices.

ANITA DEVIVO
Executive Editor

Journal Staff

ARTHUR HAUPT
Supervising Technical Editor

KENDALL C. PATTERSON
Technical Editor

ROBERT J. HAYWARD
Advertising Manager

ANNE REDMAN
Subscription Manager

PSYCHOLOGICAL REVIEW

Copyright © 1975 by the American Psychological Association, Inc.

VOL. 82, No. 4

JULY 1975

A Schema Theory of Discrete Motor Skill Learning

Richard A. Schmidt

Department of Physical Education, University of Southern California

A number of closed-loop postulations to explain motor skills learning and performance phenomena have appeared recently, but each of these views suffers from either (a) logical problems in explaining the phenomena or (b) predictions that are not supported by the empirical evidence. After these difficulties are discussed, a new theory for discrete motor learning is proposed that seems capable of explaining the existing findings. The theory is based on the notion of the schema and uses a recall memory to produce movement and a recognition memory to evaluate response correctness. Some of the predictions are mentioned, research techniques and paradigms that can be used to test the predictions are listed, and data in support of the theory are presented.

The field of motor behavior has become extremely interesting in the past few years, probably as a result of some rather severe changes in the way in which researchers have tackled their problems. Prior to about 1960, the area seemed to be dominated by the "task-oriented" approach (Pew, 1974). This approach emphasized "global" motor learning theories, such as that of Hull (1943), or emphasized no theories at all, and the area was dominated by experimenters who were testing the effects of a large number of independent variables on the overall learning and performance of motor tasks. Scoring was usually in terms of some very gross index of responding, such as time on target for 30 sec, and there was little concern for the events that changed within

the individual that enabled him to perform or to learn the motor task.

Since 1960, however, there has been a considerable shift in emphasis in motor skills research. Motor behaviorists have begun to ask questions about the kinds of processes occurring as the individual performs and learns the motor response. The tasks used have tended to shift from those that could only be scored with global measures to those that enabled the isolation of various processes and strategies and provided information about contributions of various subsystems. A number of researchers took the lead from the new directions provided by Fitts and his colleagues (e.g., Fitts, 1954) concerning the processing of information in skills. Examples of these new concerns were the time to process visual information (Keele & Posner, 1968), the development of error detection mechanisms (e.g., Schmidt & White, 1972), and the locus of attentional requirements in simple movements (Ells, 1969). Questions such as these were not popular under the earlier traditions.

Appreciation is extended to my colleagues who provided many useful comments on an earlier draft of this paper: Jack A. Adams, Sheila Henderson, Ronald G. Marteniuk, Karl M. Newell, Richard W. Pew, and H. T. A. Whiting.

Requests for reprints should be sent to Richard A. Schmidt, Department of Physical Education, University of Southern California, Los Angeles, California 90007.

A part of the reason for this shift in emphasis has been the inadequacy of the earlier theories in providing reasonable explanations for motor learning and performance; this is well documented by the controversy about Hull's (1943) theory that raged in the late 1950s and early 1960s (see Adams, 1964, for a review) and finally resulted in an almost total loss of interest in these points of view. The lack of adequate theories for guiding research in motor behavior caused motor behaviorists to reach out in other directions for explanations of motor learning. Of course, an important event in a number of fields was the arrival of the information processing and cybernetic ideas that led to closed-loop theory, and the application of these notions to psychology and motor skills was welcomed with great vigor.

Of the closed-loop theories, some were quite old, for example, Bernstein (1967) in essays originally published in 1934 and 1957, but there have been a number of widely discussed recent additions. The closed-loop models of Sokolov (1969), Anokhin (1969), and Konorski (1967), plus the work of Adams (1971) and Laszlo (1967), were most recent additions to the thinking in this area. While these theories differed a great deal from one another, they had the essential features of closed-loop theory in common: There was provision for the receipt of feedback, the feedback was checked against some reference of correctness, any discrepancy resulted in an error, and the error was subsequently corrected.

As many closed-loop theories as there were, it was surprising that there were so few serious attempts to test them against one another experimentally. It is not because the theories did not produce an interest among the workers in the field, because these ideas have been used as explanations for various experimental results by numerous authors. Rather, it is because most of these theoretical ideas are not strictly theories, but are more properly termed *models* (see Lachman, 1960, for a discussion of the distinction between models and theories). Often, such models are very easy to construct, as seemingly all one needs to propose a new model is a few appropriately labeled boxes, sup-

posedly representing processes in motor performance, with connecting arrows, and a new model is born. What has often resulted are models that are not tied to experimental data for their formation and that do not provide ways for conducting tests of them. The result is a diagram that is useful perhaps in visualizing what might be happening in skills and learning, but which has little basis in fact and cannot be verified (or, more properly, disproved) by experiment. This kind of model building has been taken far too seriously by students of motor behavior.

A notable exception to this statement was the closed-loop theory proposed in 1971 by Adams. This theory was generated in a way quite different from that of the earlier closed-loop ideas. First, the original focus was an existing body of carefully controlled basic research in motor learning—that dealing with the learning of slow, graded, linear-positioning tasks. Focusing on these data, Adams attempted to explain the various findings with the aid of closed-loop notions of error detection and correction. Most important, Adams ensured that his ideas would be testable by providing operational definitions for all of his constructs, by suggesting experimental variables that, when appropriately manipulated, should produce certain changes in particular measures, and by suggesting experimental paradigms for testing various aspects of the theory. The results of this effort were seen immediately with the appearance of approximately 20 articles testing the Adams theory in the first two years after it was introduced—already more than for all of the earlier closed-loop theorists combined. Clearly, the production of a theory that was so easily testable by researchers was appealing, and the Adams theory became a most important article for the field of motor learning in a very short time.

THE ADAMS THEORY

Adams' (1971) theory proposes that there are two states of memory, termed the *memory trace* and the *perceptual trace*. The memory trace, analogous to recall memory in verbal learning, is a "modest motor program" responsible for initiating the

movement, choosing its initial direction, and determining the earliest portions of the movement. Its strength is developed as a function of knowledge of results (KR) and practice. The perceptual trace, on the other hand, is analogous to recognition memory in verbal tasks and is responsible for guiding the limb to the correct location along the trackway. The perceptual trace is formed from the past experience with feedback from earlier responses and comes to represent the sensory consequences of the limb being at the correct endpoint. During the movement, the subject compares the incoming feedback (from the eyes, ears, proprioceptors, etc.) against the perceptual trace to determine if the limb is in the correct final location; if it is, he stops responding, and if it is not (i.e., an error is signaled from the perceptual trace), he makes a small adjustment and the comparison is made again until the limb is in the correct location. With increased exposure to feedback and KR, the perceptual trace is strengthened, and the individual becomes more accurate and confident in his responding.

Some Strengths of Adams' Theory

Adams' theory has a number of characteristics that are generally considered desirable attributes. Some of these, in each case representing improvements over early closed-loop theories, are as follows.

Concern for learning. Almost every one of the early closed-loop theories has dealt with the processes thought to occur in the performance of already-acquired skills. In addition to this concern, the Adams theory is directed toward the *learning* of novel motor tasks, which was not a consideration in the earliest work.

Reduced scope. Adams has begun modestly, attempting to limit his theoretical predictions to the body of data dealing with the learning of linear positioning. Thus, the theory is very close to the data and attempts to explain them with the theoretical constructs proposed. To be sure, the theory probably does have applicability to response classes other than linear positioning, but a limited focus provides the most solid beginnings for further theoretical developments.

Empirical support for constructs. The processes postulated in Adams' theory are all known processes, in that there is considerable evidence, either directly from motor behavior or inferred from other response classes, for each. This has not been the case for many of the earlier positions, in which processes were proposed without either behavioral or neurological support. Adams' approach has been to insist that the model postulate only those kinds of processes or mechanisms that have a reasonable probability of empirical reality.

Simplicity. The Adams theory is quite simple, postulating a minimum of hypothetical states to account for the learning of positioning tasks. Some of the earlier models are extremely complex, with many times more states postulated, and it is encouraging to think that motor learning might be explained by fewer postulations than seemed necessary before.

Some Criticisms of Adams' Theory

With all of the research and thinking that has developed since Adams' theory was published, it is not surprising that some shortcomings in the theory have appeared. Some of these result from logical difficulties with the theory as stated, and others result from recent data that do not follow the theory's predictions.

Limitations to positioning responses. Although this is not a serious criticism, some researchers are worried that positioning responses are not representative of the wide range of behavior that one would like to term *skilled*, and that the theory cannot explain the learning of other types of responses. It should be reiterated that the theory was deliberately limited in scope because of the lack of good evidence in other types of responses, but it would be a desirable goal to have a theory that predicts performance and learning in other tasks as well. Specifically, the generalization to more rapid responses (e.g., kicking and throwing) would be important, as would attention to open and closed skills or to skills in which the major goal is other than accuracy (e.g., pole vaulting). The author and his colleagues (Schmidt & White, 1972; Schmidt & Wris-

berg, 1973) have attempted to extend Adams' theory to simple accuracy tasks requiring rapid movement (approximately 30 cm in 150–200 msec), but the theory was not specifically designed to handle these tasks and tests using such tasks are not strictly tests of Adams' point of view. Some revision seems clearly necessary in order to extend the theory to more rapid movements.

The error detection mechanism. The major feature of Adams' theory is that it provided a means for the subject to determine, in the absence of KR, his error for the response just produced, to use this information (termed *subjective reinforcement*) as a means for maintaining performance, or even to continue to learn without KR. The author has argued before (Schmidt, 1974, Schmidt, Note 1) that learning without KR does not follow logically from Adams' theory. Theoretically, the subject uses the perceptual trace and feedback during the movement to guide the limb to the proper location, and he moves to that position he recognizes as correct; thus, the movement endpoint is that position for which the error signal (the difference between the perceptual trace and incoming feedback) is zero. After the subject has removed his hand from the lever, how can he generate an additional error signal that is sensitive to the difference between the position on the trial and the correct position? The answer is that he cannot. When asked what the error was on that trial, the subject can provide a guess, but that guess is probably uncorrelated with the actual error for the trial, because his error signal at this point is necessarily zero. Thus, it would seem that the theory cannot provide subjective reinforcement that would enable the individual to continue to learn without KR.

In a study by Schmidt and Russell (unpublished study, 1972),¹ subjects were given 100 positioning trials, with KR after each trial. A paradigm was used in which the subject responded, was asked what he

thought his error was (in the units of the positioning task), and was then provided KR. The measure of the error detection mechanism was the correlation, computed within subjects, between the actual and judged error on the last 20 trials. The highest correlation was .40, and many of the r s were negative, with the average being .21. Had there been a strong error detection mechanism operating after the movement, larger correlations, similar to those found by Schmidt and White (1972) for a rapid movement task ($r = .90$), should have been found. The Schmidt and Russell findings agree with the prediction that there should be no error detection mechanism after the movement in slow positioning.

Learning without KR. Another important prediction of the Adams theory is that later in learning, when the perceptual trace is well established, the subject should be able to continue to learn without KR. There is no evidence that individuals can continue to learn positioning tasks after KR withdrawal, and the best approximation seems to be that they can maintain performance reasonably well after KR withdrawal (e.g., Bilodeau & Bilodeau, 1958). But in spite of the lack of evidence, the theory cannot predict this no-KR learning. If, on the first no-KR trial, the movement is not perfect, the perceptual trace will be degraded somewhat because of the feedback from that response. The next response will be less accurate because the perceptual trace has been slightly weakened, and the perceptual trace will be degraded further, and so on. Eventually, the perceptual trace becomes increasingly weak, with a corresponding decrement in accuracy, and performance is increasingly inaccurate. Clearly, the theory cannot predict learning without KR for positioning responses.

Generalization to rapid responses. While the Adams theory was not intended to be explanatory for rapid responses, a number of experimenters (e.g., Newell, 1974; Newell & Chew, 1974; Schmidt & White, 1972; Schmidt & Wrisberg, 1973) have used Adams' closed-loop ideas to generate predictions for this response class. The reasoning was that in slow positioning responses the memory trace and perceptual trace become

¹ Requests for information concerning this and other unpublished studies cited in this article should be sent to Richard A. Schmidt, Department of Physical Education, University of Southern California, Los Angeles, California 90007.

experimentally confounded, with positioning tasks not providing measures that could be ascribed to the variations in the memory-trace strength. One solution was to use rapid movements, usually the movement of a slide for 10–15 in. (25.4–38.1 cm) with a goal of 150–200 msec, where the movement would be carried out by the memory trace (in this case a motor program without feedback involvement), with the perceptual trace strength evidenced by how accurately the subject could guess his movement time after the movement.

Many of the predictions that seemed to come from Adams' theory were supported. The strength of the error detection mechanism (the perceptual trace measured as the within-subject correlation between actual and judged error) increased as a result of practice (Schmidt & White, 1972), the error detection mechanism was sensitive to the amount of experience with feedback stimuli (visual and auditory) present in the task (Newell & Chew, 1974; Schmidt & Wrisberg, 1973), and ratings of the subject's confidence increased with practice as well. But there were some instances of nonsupport. Shea (unpublished study, 1972) failed to confirm that the interresponse interval was a factor in development of either the memory trace or the perceptual trace. Also, Adams states clearly that the delay of KR should not be a factor in the development of the perceptual trace, and yet a study by Schmidt, Christenson, and Rogers (in press) indicates that delaying KR by 25 sec versus 5 sec caused a drop in the correlation (from .92 to .61) between actual and judged error; the interpretation was that KR delay retarded the development of the perceptual trace.

Perceptual trace development. Adams' theory holds that the perceptual trace is formed from feedback traces associated with having moved to the correct location, and that without this experience at the correct location, the perceptual trace cannot develop. Williams and Rodney (in press) had subjects practice a linear positioning task under a variety of conditions, and two of these conditions test this prediction. Group TO (target only) moved 16 times to a stop that

defined the criterion location, while Group IR (interpolated random) moved to 16 randomly ordered stops (not at the criterion location), after being told that the criterion location was in the center of these positions. When subjects then attempted to move to the criterion position without the stop for 20 trials without KR, the absolute errors (see Figure 1) for the two groups were nearly equal on the first block of 4 trials; also Group IR maintained performance over blocks, whereas Group TO regressed significantly. The finding that subjects who had never experienced the correct location (Group IR) could move as accurately (and maintain accuracy more effectively) as those subjects who had the conditions necessary for the development of the perceptual trace (Group TO) is very damaging to Adams' position.

The storage problem. Most existing theories, whether they stress the open-loop (programmed) aspects (e.g., Henry, 1960) or the closed-loop aspects of movement control (e.g., Adams, 1971), implicitly postulate that for each movement that is to be made there must be either a motor program or a reference against which to compare feedback (depending upon the type of theory), and that there is a one-to-one mapping between stored states (either programs or feedback states) and movements to be made. This presents problems for the central nervous system in terms of the amount of material that must be stored, as discussed recently by MacNeilage and MacNeilage (1973) in the speech production area and by Schmidt (in press) for motor skills. They estimated that for the English language, considering inflections and accents, there were some 100,000 phonemes required, and thus the same number of stored states. When we add to this the nearly countless additional ways (other than speech) in which individuals move their musculature, the individual must have a nearly countless supply of either programs or feedback states in storage. While this may be possible—and there is no evidence that it is not possible—it would seem desirable to postulate mechanisms that do not require this level of storage. Recently, a number of attempts at

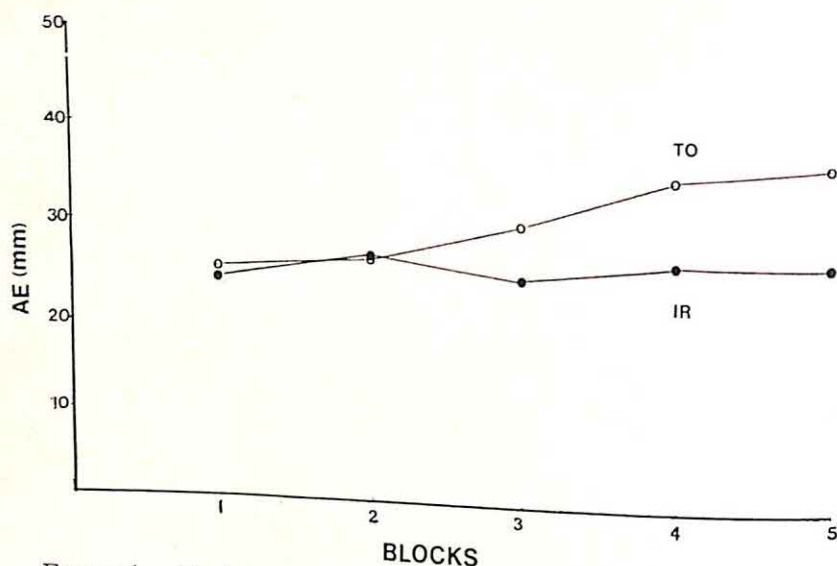


FIGURE 1. Absolute error (AE) in linear positioning as a function of variability of prior experience (from Williams & Rodney, in press). Abbreviations: TO = target only group; IR = interpolated random group.

alternatives to the earlier program or feedback theories have been proposed that do not require this volume of material to be stored, such as the MacNeilage (1970) target hypothesis and the use of reflexes by Easton (1972).

The novelty problem. A companion to the problem of one-to-one storage of programs or feedback states is the problem of how the performer produces a "novel" movement. It has seemed clear to investigators for many years—for example, Bartlett (1932) and Bernstein (1967) in essays originally published in 1934 and 1957—that when we make a motor response in a game, for example, we do not execute the movement exactly as we have made it before, and this is borne out by recent biomechanical analyses of movements (e.g., Higgins & Spaeth, 1972). If the response is to be programmed, for example, the sequence of muscle commands would be appropriate for only one movement, beginning with the body in a specific position, and with an identical goal; and it is probably true that the same response is never made twice when one considers the number of possibilities there are, for example, in shooting a basketball. Therefore, motor-learning theory must be able to account for the generation of such novel movements to overcome the shortcomings of existing theories in this regard.

THE GOAL OF THIS ARTICLE

The research and thinking that Adams' (1971) theory, and other closed-loop theories as well, have generated in the last few years has revealed a number of shortcomings—both logical and empirical—as outlined in the previous sections. The goal of the present article is a theory that contains the strong points of the Adams view but is considerably different in a number of important ways, so that it will be consonant with the experimental literature. The theory attempts to deal with discrete tasks, that is, those that have a recognizable beginning and end and are usually quite short in time (e.g., less than 5 sec in duration). A number of types of discrete tasks are considered. In addition to the linear positioning responses with which Adams dealt, the theory is aimed at rapid "ballistic" tasks with very short movement times (e.g., less than 200 msec), tasks that are "open" as well as "closed" (Poulton, 1957), tasks that demand accuracy, and tasks that have maximum speed, height, etc., as goals. Tracking and other continuous tasks are deliberately excluded from this article because of the rather complex simultaneous interaction between sensory input and motor output, although there is the possibility that the ideas presented here will be useful at some future time for tracking

(see Pew, 1974, for an up-to-date review of the tracking literature).

THE THEORY

In order to correct for the shortcomings of existing open- and closed-loop theories in accounting for the recent motor learning and performance data, a somewhat marked departure from these original points of view seems essential. As with any new theory, there is probably not very much really new, with many of the ideas borrowed from earlier points of view. Such is the case here. As the theory is explained in the next few sections, it will become obvious that the lineage of the major ideas can be traced to Bartlett (1932) in terms of the notion of the schema, to Adams (1971) for his application of closed-loop theory to learning of motor skills, to Pew (1974) for the suggestions about the application of the schema to motor skills, and to Lashley (1917) for his lead in characterizing man as controlling his movements centrally with "motor programs." It is to the open-loop control ideas and the motor program that we turn next.

The Motor Program

The idea that the human being has a set of stored muscle commands ready for action at any time has probably been with us for a very long time, but the first important documentation that movement was centrally controlled was provided by Lashley (1917) in describing the movements of a patient with a gunshot wound in the back. Because of the wound, the patient had lost all sensation from his lower limbs, but had not lost the efferent pathways that enabled him to move. Even though he could not feel movement in his leg, he was nevertheless able to position it with surprising accuracy, not unlike a normal control subject. This finding led Lashley to argue for a position in which movement was controlled centrally, since there was little possibility that the wounded patient could have been using feedback to guide his movements.

This idea has been restated many times, both formally and informally, and with the advent of the electronic computer the centralist notion became couched in terms of

the programs used by the machines; Henry's (1960) "memory-drum theory" was such an idea, using the notion of the memory drum element in the earlier computers as an analogy to the human system. One of the more recent statements is that of Keele (1968), who defined the motor program as a sequence of stored commands that is "structured before the movement begins and allows the entire sequence to be carried uninfluenced by peripheral feedback" (p. 387).

The primary evidence for the motor program notion has been that the processes involving the generation of sensory error information, perceiving it, and initiating corrections in response to those errors was quite slow, requiring from 120–200 msec (about one "reaction time") to initiate the corrections. While it is true that there is some variance in the speed with which the various sensory channels operate, with proprioception being fastest, at about 110 msec (Chernikoff & Taylor, 1952), and vision being the slowest, at about 190 msec (Keele & Posner, 1968), there is still the problem that many movements can be carried out in far less time than is required for the feedback loop to operate. In fact, the strongest human evidence for the motor program notion seems to be that subjects can initiate, carry out, and stop a limb movement within 100 msec, implying that decisions about when to stop the movement must have been made prior to the initiation of the movement. These results have provided a serious dilemma for the closed-loop performance theorists, and have provided the support needed for the program notion to become popular.

The idea of the motor program is largely a default argument, as pointed out by Pew (1974). There is really no direct human evidence of a motor program; centralists reason that there is no other known means of producing the movements; thus programs must be the explanation. Actually, it should be shown either (a) that feedback is present in movement but is not used, or (b) that feedback is not present and movement can still occur. Strictly, neither of these two possibilities has been shown experimentally.

There are, however, data from subhuman species that provide support for the program

notion. The most convincing is the report from Wilson (1961), who totally deafferented the wing musculature and related joints of locusts. When a ganglion near the head was stimulated electrically, the locust would produce movements of the wings closely resembling flying (although the movements were decreased in amplitude somewhat), and these movements would continue for an extended period of time without further stimulation. Thus it appears that there are motor programs, and that they are capable of carrying out movement in the absence of sensory feedback. Similar data have been collected by Nottebohm (1970) with birdsongs. The limitation of these findings for human motor behavior is that these subhuman motor programs can probably be considered innate, and as yet there are no data showing that learned acts can be programmed, as we would like to believe in terms of learning a program for kicking a football. Even so, the notion seems to have sufficient direct and indirect support to warrant its inclusion here.

The motor program notion is modified slightly in the present situation, however. The original form of the motor program implies that every movement must have a separate motor program associated with it, and such a postulation brings in the storage problem mentioned earlier. The notion has therefore been changed somewhat to mean that there are generalized motor programs for a given class of movement. For example, there might be a single program for the many ways of throwing a baseball. How wide this category of movements governed by a single program might be has been the subject of some debate, as some would hold that the movement category is quite narrow (e.g., Henry, 1960), while others, referring to the "overarm pattern," imply that the program might govern all movements in which something is propelled overhead (e.g., Broer, 1973). How big the category actually is does not matter a great deal, however, as long as it can be postulated that there is not a one-to-one match between the program and each specific movement that the individual can produce, in order to avoid the storage problem.

These generalized motor programs are assumed to be able to present the prestructured commands for a number of movements if specific response specifications are provided. Thus a motor program for throwing a ball could be modified by specific instructions to throw fast or slow. These specifications can be thought of as parameters that can be varied before the movement begins to enable the execution of the program at a different speed, a different force, and so on. Thus, the performer's problem in choosing a movement is the determination of the response specifications that will modify the existing stored motor programs.

According to the current literature in the neurological control of movement (e.g., Granit, 1970), the output from the motor program consists of two sets of signals. The alpha efferent signals innervate the extrafusal fibers in the main body of the musculature, while the gamma efferent signals innervate the intrafusal fibers in the muscle spindles. Through the servo-action of the muscle spindles, the alpha and gamma activity is coordinated, so that minute variations from the intended spatial-temporal pattern of movement are corrected very rapidly—within 30–50 msec—via the rapid, spindle-initiated feedback loops. These servo-type corrections control for unplanned variation in the fatigue state, unexpected forces occurring in the limbs, etc., so that the movement takes on the intended path (i.e., the path intended when the program was chosen). Thus, strictly speaking, it makes little sense in view of the evidence to claim, as Keele (1968) and Schmidt (1972b) have done, that movements with durations less than 200 msec take place "without involvement from peripheral feedback," because the spindle system operates much more quickly than this. What is meant by this definition of the motor program is that when the program is initiated, it carries itself out as planned, correcting for deviations from the intended path of the movement, but that if something happens in the environment that requires that some new movement (a new goal) be planned, the performer cannot accomplish any such changes until the program has run its course for approximately

200 msec. Thus, control is open loop because stimuli from the periphery cannot initiate a new program until the present one has run its course for one reaction time (see Schmidt, in press).

Actually, the time over which a program must run its course without change is probably far longer than the lower limit of 200 msec usually specified. If the movement time is only 200 msec, any signal that the subsequent movement is about to be incompatible with the environmental conditions must occur very early in the movement; being able to correct for errors occurring this early seems unlikely, because the initial portions of the movement must usually be made in order for the subject to perceive that it is going to be incorrect. Depending upon the type of response, it is reasonable to believe that the program could carry itself out for nearly 400 msec before it could be changed. The amount of possible feedback involvement in a movement longer than one reaction time is probably dependent primarily upon the movement time, and seems independent of the movement velocity so long as the movement time is constant (Schmidt & Russell, 1972).

A primary focus of the present theory is on the development of motor programs and their response specifications, an emphasis shared by a number of motor skills researchers. Perhaps this point is made best in a passage from a recent article by MacNeilage and MacNeilage (1973):

The need for peripheral sensory feedback can be thought of as inversely proportional to the ability of the central nervous system to predictively determine . . . every essential aspect of the following acts. (p. 424)

Because of the lags in processing feedback, the subject becomes less and less dependent upon feedback for performance, and the emphasis shifts from feedback-controlled, jerky performances to the smooth execution of almost completely open-loop movements. Thus, the problem for the subject in learning motor skills is to develop these open-loop programs for his movements to free himself from feedback involvement. Pew (1966) has shown, in a task requiring alternate finger-tapping movements to keep a dot cen-

tered on a screen, that there was a shift in control from closed-loop to open-loop control. Schmidt and McCabe (Note 2) also found less feedback involvement in a discrete timing task with a 750-msec movement time, using the index of preprogramming discussed by Schmidt (1972b).

The Schema

The notion of the schema is not a new idea at all, as the first formal statement of the idea was made by Head in 1926, and the ideas were subsequently modified considerably and presented in a book by Bartlett called *Remembering* (1932). The idea, usually stated with respect to perception, is that in order to perceive a set of visual stimuli (e.g., a dog) and to classify these stimuli correctly in the category "dog," we need not have previously received the particular set of stimuli in question. Through our past experiences with seeing dogs, we store these stimuli in recognition memory and also abstract these stimuli into a concept related to dogs for additional storage. This concept forms the basis of a *schema* or rule for determining whether a new set of visual stimuli should be classified into the category "dog" or not. Thus to recognize an animal as a dog, we need not have ever seen that particular animal before, and with the use of the schema for dogs, we correctly identify the animal's category. This rule for determining the category membership of a set of stimuli forms the basis for the definition of the schema, as defined by Evans (1967b):

A schema is a characteristic of some population of objects, and consists of a set of *rules* serving as instructions for producing a population prototype (the concept). (p. 87)

Thus, given the schema for "dog," which consists of a set of rules for determining if a set of stimuli should be classed as a dog or not, the individual comes to the decision about the category membership of the stimuli.

While the notion of the schema has been in existence for a long time, it has been limited in its usefulness theoretically because there have been a number of ways of thinking about the same idea (e.g., Head, 1926,

versus Bartlett, 1932), the ideas were somewhat too "mentalist" for the behavioristic leanings of psychology until recently, and there have not been attempts to operationalize the concept so that it could be tested experimentally. Recently, however, a number of researchers have invoked the notion to explain nonsense pattern recognition, and Evans and his colleagues (Edmonds & Evans, 1966; Edmonds, Evans, & Mueller, 1966; Edmonds & Mueller, 1966; Edmonds, Mueller, & Evans, 1966; Evans, 1967a, 1967b; Evans & Edmonds, 1966) and Posner and Keele (1968, 1970) have provided strong operational definitions and tests of the concept. The descriptions of the Posner and Keele studies indicate how the schema has been tested for pattern recognition tasks.

Posner and Keele (1968, Experiment III) presented a series of random dot patterns on a screen to subjects. They had three basic 9-dot patterns (the "prototypes"), and they produced variations (called "distortions") of these prototypes by randomly moving the dots in the original patterns. In training, 24 of the distortions were presented, and the subjects learned with practice and knowledge of results (KR) to categorize them correctly (Edmonds, Mueller, & Evans, 1966, have shown that KR is not necessary for this categorization to occur). The original prototypes from which the distortions were created were not shown. Following training, a transfer test was given, consisting of the 3 original prototypes (not seen previously), 6 distortions shown in training, 12 distortions not previously shown, and 3 new, unrelated, random patterns.

On the transfer test subjects classified the distortions they had seen previously most accurately (13.0% error), but they correctly classified the prototype patterns nearly as well (14.9% error). Classification of the distortions not previously seen was not as accurate (26.9% and 38.3% error, depending upon the degree of distortion). These data indicate that subjects can correctly classify dot patterns that they have never seen before, implying that they generated an abstraction (a schema) of the dot patterns of a given category, stored it, and were able to use this

stored abstraction to recognize the prototype pattern when it was later presented. A subsequent study by Posner and Keele (1970) showed that when the transfer test was provided 1 week after the original learning, the recognition of the prototype did not decrease over the retention interval, whereas the recognition of previously seen distortions of it did. This implies that the schema for the dot-pattern classes was retained nearly perfectly, while the actual patterns seen previously could not be stored as effectively. These data clearly suggest that subjects store both the patterns seen and the schema (the abstraction) of the pattern (although at different strengths), and that the schema allows subjects to recognize the prototype pattern without ever having seen it previously. These data, in addition to the Evans experiments cited previously, provide an experimental paradigm and rather strong evidence for the schema concept, at least for stimulus recognition. The next sections deal with the extension of the schema to the area of motor response production and motor response recognition, respectively.

The Motor Response Schema

Bartlett (1932) discussed the notion of the schema as a means of solving the storage problem for response production and as a means of generating novel responses. He was quite clear that some such mechanism must exist, as evidenced by the following:

How I make the [tennis] stroke depends upon the relating of certain new experiences, most of them visual, to other immediately preceding experiences and to my posture, or to balance of postures, at the moment. (p. 201)

Then, after one takes in this information about the present state of the body and environment, Bartlett says,

When I make the stroke I do not, as a matter of fact, produce something absolutely new, and I never merely repeat something old. (p. 202)

He was not very clear concerning precisely how the schema would operate and how learning would be handled, but recently Pew (1974) has provided some additional thinking about the schema. While the views expressed in this article are not exactly like

Pew's, they clearly had their origin in Pew's thinking.

The notion of the schema as an abstraction of a set of stimuli requires some modification to allow application to response production. Basically, when the individual makes a movement that attempts to satisfy some goal, he stores four things: (a) the initial conditions, (b) the response specifications for the motor program, (c) the sensory consequences of the response produced, and (d) the outcome of that movement. The next section explains each of these constructs.

Initial conditions. A number of writers (e.g., Keele, 1968; Pew, 1974) have indicated that in order for the subject to move effectively, he requires information about the preresponse state of his muscular system and of the environment in which he is to move. The initial conditions, then, consist of the information received from the various receptors prior to the response, such as proprioceptive information about the positions of the limbs and body in space, as well as visual and auditory information about the state of the environment. After the movement, the initial conditions used to plan the movement are stored.

Response specifications. Since the motor program for generation of the muscle commands is assumed to be rather general, with variations of the basic pattern possible by changing such important elements as the speed with which it is run off, the forces involved, etc., the subject must specify these elements before the movement can be run off. After the movement, these specifications are stored along with the other information received after the movement. These serve as a record of the specifications of the movement produced.

Sensory consequences. The third type of information stored after the movement is the response-produced sensory information. This information consists of the actual feedback stimuli received from the eyes, ears, proprioceptors, etc. Thus, the sensory consequences are an exact copy of the afferent information provided on the response.

Response outcome. The fourth source of information stored after the movement is the

success of the response in relation to the outcome originally intended. The desired outcome (or goal) of the movement is potentially a verbalization, such as, "throw the dart very hard at the center of the target" (although the desired outcome need not actually be verbalized), and the response outcome is in these same terms, such as, "you threw 23 mm to the left." Thus, the actual outcome of the movement is stored, not what was intended. This outcome information arises from information the subject receives after the movement, and consists of KR (when present) and subjective reinforcement that the subject obtains from other sources of feedback. The accuracy of the outcome information is thus a direct function of the amount and fidelity of the feedback information, and a subject without any feedback information does not have outcome information to store.

Schema formation. The above four sources of information—initial conditions, response specifications, sensory consequences, and response outcome—are stored together after the movement is produced. When a number of such movements have been made, the subject begins to abstract the information about the relationship among these four sources of information in a way suggested by the dot-pattern discrimination experiments. The schema notion requires some extension from the original pattern-perception idea, however, in that in the motor case it is the relationship among the arrays of information that is abstracted rather than the commonalities among the elements of a single array. The strength of the relationship among the four stored elements increases with each successive movement of the same general type and increases with increased accuracy of feedback information from the response outcome. This relationship is the schema for the movement type under consideration and is more important to the subject than is any one of the stored instances, which, according to Posner and Keele (1970), are forgotten more quickly over time than is the schema. Figure 2 diagrammatically shows how these sources of information are associated to form the schemata.

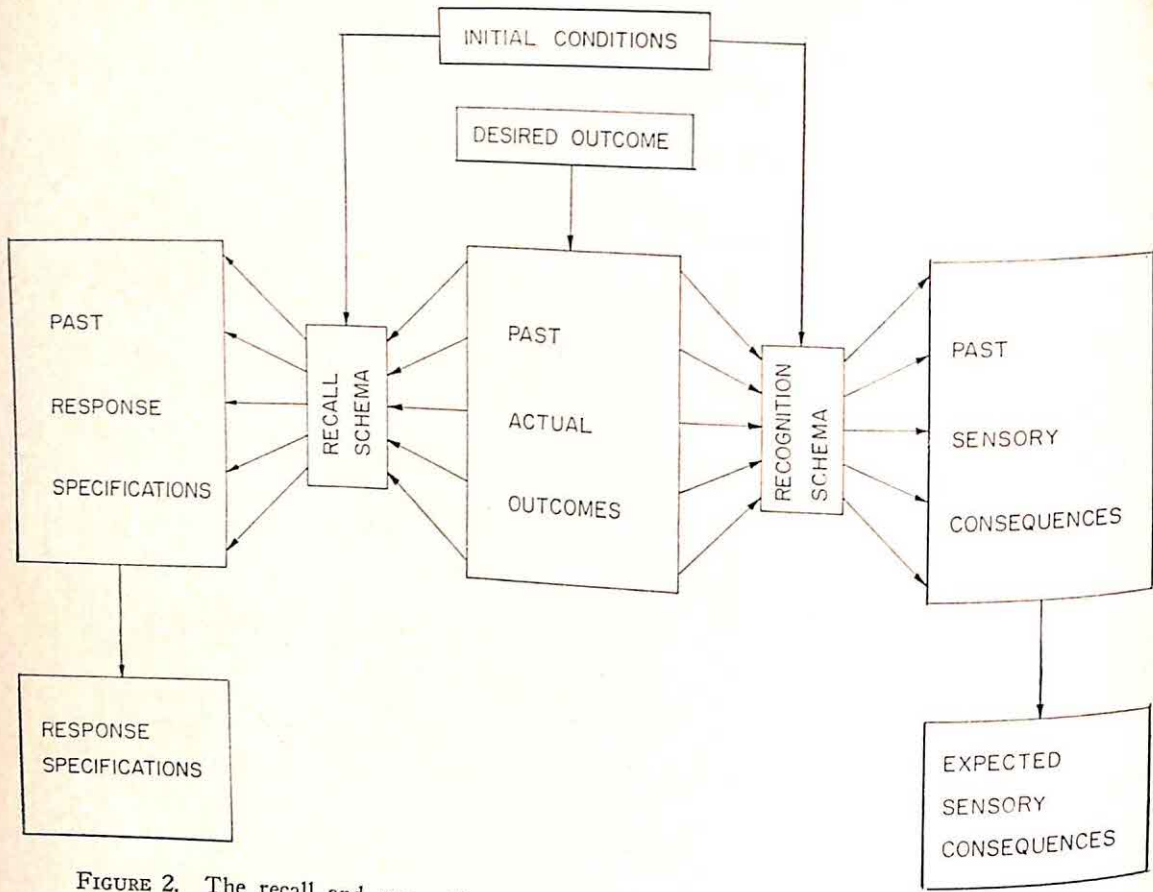


FIGURE 2. The recall and recognition schema in relation to various sources of information.

Response production. When an individual is required to make a response of a type for which he has a schema already developed, he begins with two inputs to the schema: the desired outcome for the movement and the initial conditions. From the relationship between the past outcomes and response specifications (the recall schema), he determines what set of specifications will achieve the desired outcome. The subject need never have produced those specifications previously, because they are determined from a combination of initial conditions and an outcome that might never have been present earlier; nevertheless, with the schema rule between the outcome and response specifications, as modified by the initial conditions, the specifications can be determined as interpolations among past specifications. When the specifications are determined, the subject executes the motor program with the particular set of specifications, and the movement is carried out. Be-

cause the specifications may have never been used in exactly this way before, the movement that results may be novel, in that it may, strictly speaking, never have been executed before.

Response recognition. At the same time that the subject uses the schema to generate the response specifications, he also generates the expected sensory consequences of the movement. Because the schema also contains a relationship between the past outcomes and the past sensory consequences, given the desired outcome, the subject can generate two types of sensory consequences, modified by the particular initial conditions. The first expected sensory consequence is the expected proprioceptive feedback, which should result if the desired outcome is achieved, and which consists of the anticipated feedback from the various proprioceptors in the muscles and joints, as well as the anticipated information from the vestibular apparatus. Second, the anticipated ex-

teroceptive feedback consists of anticipated vision, audition, etc., of the environment and of the limb and the objects moved by it.

During and/or after the movement, each of these expected sensory consequences is compared with the respective inflow of sensory information (the proprioceptive and exteroceptive feedback, respectively), and a resulting mismatch in the expected and actual sensory consequences produces an error that is fed back to the schema, providing information (subjective reinforcement) as to the outcome of the response produced. These processes are symbolized in Figure 3, in which the recall and recognition schemata are combined (i.e., the "motor response schema") to increase the clarity of presentation.

Since the expected sensory consequences are dependent upon the developing relationship between the response outcome (as determined by KR or other subjective reinforcement) on previous trials and the sensory feedback actually received, the strength of the recognition memory should increase as a function of both KR in initial practice and the quality and amount of feedback received on each trial. This portion of the schema rule is the basis of recognition memory for the movements of the type governed by the schema. This recognition memory is assumed to be analogous to recognition memory in verbal learning—including the assumption that it is independent of motor recall memory (i.e., that memory associated with the response-production portion of the schema), even if some of the variables (e.g., KR and subjective reinforcement) are similar for the two memory states.

An important point should now be made about the nature of the expected sensory consequences generated before the movement. In order for the subject to be able to receive information about the correctness of the movement in relation to the desired outcome, he must be able to compare the actual feedback with the feedback expected if the movement achieved the environmental goal, and thus the expected sensory consequences must represent the feedback consequences of this correct movement. This is in contrast to a number of earlier theories

(e.g., Anokhin, 1969; Bernstein, 1967; Pew, 1974; Sokolov, 1969) in which the expected sensory consequences are of the movement actually chosen; such expected sensory consequences can provide information about the extent to which that movement was carried out faithfully, but can provide no information about whether appropriate response specifications were selected. A major difference between these theories and the present one concerns the way in which these expected sensory consequences are chosen.

In the present theory, the subject begins the process by selecting a desired outcome and by noting the initial conditions at the time. The relationship between actual outcomes and sensory consequences (the recognition schema) allows the generation of a set of expected sensory consequences that represent the best estimate of the sensory consequences of the correct movement. Also, the relationship between actual outcomes and response specifications (the recall schema) permits the generation of the response specifications that are the best guess as to how to achieve the desired outcome. With this method of selection, even though the expected sensory consequences and response specifications are strongly associated, there is no necessity that they be isomorphic, as with the earlier theories, because the recall and recognition schemata are separate. While the two schemata do share the initial conditions and actual outcomes as variables, they are separate because the recall schema is the relationship between these variables and response specifications, whereas the recognition schema is the relationship between these two variables and sensory consequences. Further, it is possible that the subject may choose inappropriate response specifications on trial n , recognize his error, update his recall schema, and produce different specifications on trial $n + 1$, even if the initial conditions are identical; even though two sets of response specifications were generated on the two trials, the expected sensory consequences could have remained constant because the desired outcome did.

In the earlier theories, however, the subject first chose the particular movement pattern and then chose the expected sensory

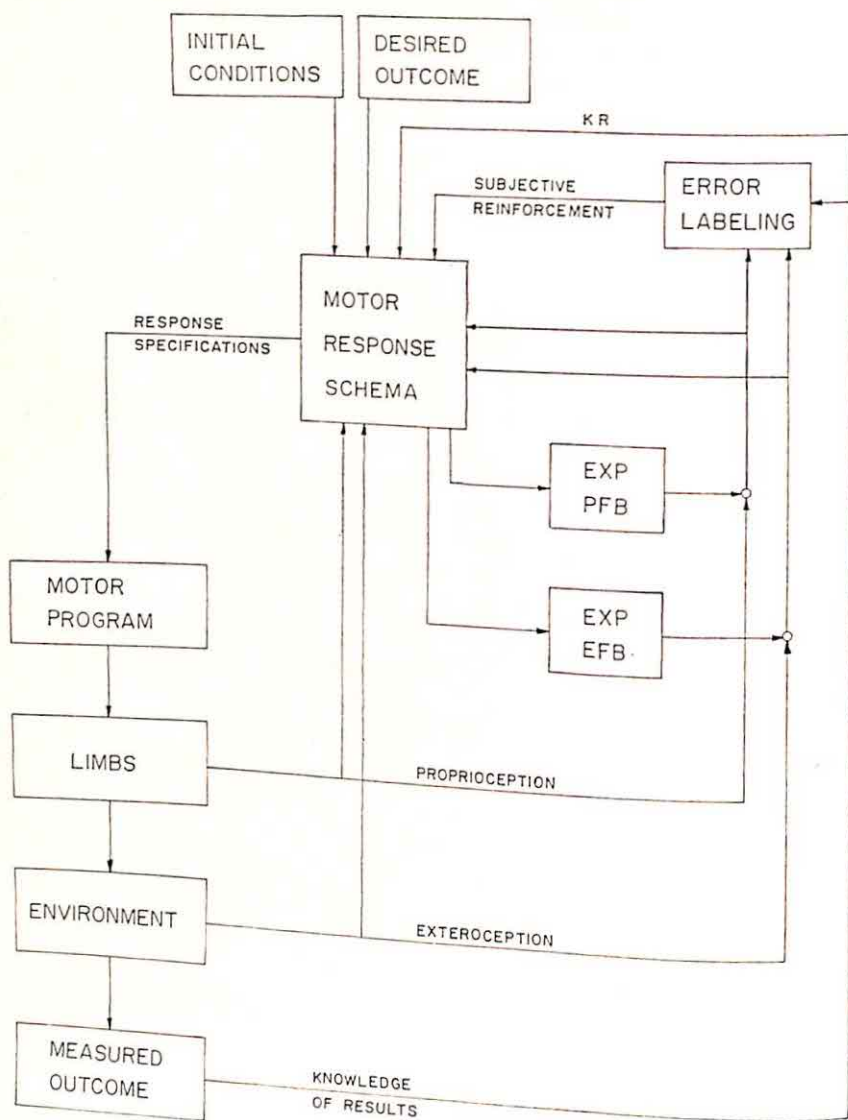


FIGURE 3. The motor response schema in relation to events occurring within a trial (recall and recognition schemata are combined for clarity). Abbreviations: KR = knowledge of results; EXP PFB = expected proprioceptive feedback; EXP EFB = expected exteroceptive feedback.

consequences associated with that response; here the expected consequences and response specifications are necessarily isomorphic because one was specified first and the other was chosen as a function of it. The limitation in this system is that if the recall schema is weak and the subject chooses inappropriate response specifications, the movement will be incorrect, but he will have no way of detecting his error, since the expected and actual sensory consequences, being of the movement chosen, necessarily match. Adams

(1971) successfully avoided this problem in his theory by insisting that the perceptual trace (the expected sensory consequences) be developed through past experience with the correct movement.

The Labeling of Error Signals

The present theory proposes that the error signal generated by the comparison of actual and expected sensory consequences exists in two possible states: (a) the raw sensory signal that arises from the compar-

ison of expected and actual consequences, and (b) this raw signal after it has been labeled and converted into a reportable form. This notion is certainly not new, as Bartlett in 1932 said,

Perceptual processes, in fact, involve two different, but related functions: (a) that of the sensory pattern, which provides a physiological basis for perceiving; and (b) that of another factor which constructs the sensory pattern into something having a significance which goes beyond its immediate sensory character. (p. 188)

While this raw signal can be used by the subject, most notably in those situations where the subject makes a series of adjustments in order to reduce his error signal to zero in positioning tasks, it is the labeled error signal that plays a role in explaining learning. The labeled error signal is termed *subjective reinforcement* and can serve as a substitute for knowledge of results (KR), providing outcome information that can update the recall schema. Subjective reinforcement is less accurate than is the perfect KR, and so the subject will probably use KR to update the recall schema when KR is present, but when KR is not present, the subject can resort to the less accurate subjective reinforcement for his outcome information.

The error labeling system is proposed to be another schema, in this case a schema for labeling sensory signals. It is assumed that past sensory signals have been stored along with the actual sensory consequences (based on KR), and a schema rule is built up over time that relates the KR received to the signals received. When this schema is well established, it enables the subject to receive error signals from the expected-actual comparisons, and to attach a label to these signals according to the KR-sensory-signal relationship. Thus, in later practice, the subject can attach a label to a new sensory error signal that he has not experienced previously, and the result is subjective reinforcement. Of course, the essential ingredient for the error labeling system is KR, and without it subjects cannot develop the schema for labeling new errors.

Producing a Movement

Given the notions presented in the previous sections, it will be useful to describe how the schema operates in producing a motor response by explaining the steps in the movement in the order that they are thought to occur. Figure 3 shows the movement process, beginning with the specification of a desired outcome and subsequent determination of the initial conditions. From the relationship established in past responses between outcomes and response specifications, new specifications for the motor program are selected. Also, at the same time, the relationship between outcome and sensory consequences allows the selection of the expected proprioceptive feedback and exteroceptive feedback. When these processes have occurred, the movement can be initiated by running off the motor program.

Immediately following the initiation of the motor program, the impulses begin to flow out to the muscular system with all of the details of at least the first 200 msec of the movement specified. As the movement is carried out, sensory receptors in the body provide information about the movements occurring. For convenience only, these receptors are classed into two groups—proprioceptive feedback and exteroceptive feedback—although it should be recognized that this classification does not represent any meaningful differences in the way these feedback sources are assumed to operate. The feedback information is fed back to the expected proprioceptive and exteroceptive feedback states, respectively, and the discrepancy between these anticipated and actual states represents an error in responding. The raw error is fed back to the schema so that in positioning movements, for example, additional adjustments can be made to reduce the error to zero. This error is also fed to the error labeling system, where the subject assigns a reportable label to it, and this resulting subjective reinforcement is then fed back to the schema as subjective information.

The final source of error is the information that the experimenter or teacher provides for the subject after the movement, termed KR. It arises from the measured

outcome of the response ("the ball missed the target by 20 cm to the left"), and is already in a form that is interpretable as a deviation from the desired outcome. The KR information, in addition to being fed back to the schema for updating the schema rules, is fed to the error labeling system to enable this system to improve its accuracy in labeling future error signals arising from the deviations of proprioceptive and exteroceptive feedback from their respective anticipated states.

Learning is possible by feeding back the essential error information to the schemata. The response specifications and initial conditions are stored when the movement is selected, and the actual proprioceptive and exteroceptive feedback are stored as the movement is progressing and as these sources of information are generated. Finally, the actual outcome is stored, based on KR when it is present, but based on subjective reinforcement if KR is not present. These sources of information can then be used to update the schema rules and provide revised estimates of the expected sensory consequences and response specifications on the next trial. The traces representing these sources of information are hypothesized to be relatively weak, so that rapid forgetting of them can be expected over time. However, the traces are assumed to be held in store sufficiently long so that they can be used to update the appropriate schemata. The schemata, on the other hand, are assumed to be stored far more permanently, although some forgetting of them over time could be expected.

SOME DISTINGUISHING FEATURES

This section concerns some of the important ways in which the present theory differs from earlier points of view. In this discussion, the ways in which the theory can deal with the problems and shortcomings of the other theories are discussed.

Applicability to Various Response Types

The present theory extends the notions of Adams' (1971) theory to response types other than linear positioning, and can handle the learning and performance of discrete

tasks that are either open or closed (Poulton, 1957) and either rapid or slow. In addition, a type of discrete response not frequently discussed is one in which the performer continually strives to do "more of something," such as moving faster, jumping higher, or producing more force, and this type of task does not have accuracy, in the traditional sense at least, as a goal. Further, the theory has the capability of explaining novel movements that may not have been produced previously.

Open versus closed skills. Poulton (1957) defined open skills as those in which environmental and situational characteristics can change as the subject plans or performs his response, such as the response that would be required of a wrestler attempting a take-down. Closed skills, at the other extreme, are those in which the goal and environmental conditions are relatively constant. The important distinction between these two types of movement was recognized by Gentile (1972). Essentially, with closed skills the performer need only learn one movement that satisfies the goal, whereas with open skills, the environmental conditions are never constant, and the subject must plan his response to meet the anticipated situational demands of the environment. The question of the novel movement is particularly important for these open skills, since performers must make movements never before made if the environment is not exactly the same as for previous movements.

Theoretically, with closed skills the subject develops the schema rule so that he makes finer and finer predictions on the basis of the rule, and the expected sensory consequences are more and more accurate estimates of the consequences of that one movement. This is much the same as Adams has it (but the mechanisms are different, of course). The schema rule produces increasingly accurate estimates of response specifications and expected sensory consequences, and the movements become more and more accurate as a result. With open skills, however, the problem is similar, but it is complicated by the fact that the subject must be able to produce novel responses. He does this by again strengthening the schema rule

between outcomes and response specifications and between outcomes and sensory consequences, as mediated by the initial conditions. When a new environmental condition arises, the subject can generate the response specifications and the expected sensory consequences on the basis of his perceptions about the state of the environment (and what that state will be when the movement is actually carried out).

The present point of view, contrary to Gentile (1972), does not see open and closed skills as being fundamentally different. Closed skills are those movements in which the subject does not have the problem of environmental uncertainty as he plans the movement. Open skills, however, can be regarded as closed skills with environmental uncertainty added. The problem for the performer in these open tasks is to determine what the environment will be like when the movement is finally executed, and to plan the movement accordingly. However, because of the reaction-time lags in the motor system, there comes a time in every open skill where the individual must execute the program associated with the best estimate of the changing environment, and at this point the movement becomes "closed" for the subject. That is, given his temporal limitations, he executes the movement and cannot change it for at least 200 msec, and it is exactly as if the environment were fixed in that state predicted by the subject when he planned the movement. If the subject was incorrect in his estimates of the environmental state, he would, of course, have produced a response that was inappropriate given the environment, and he would have to initiate corrections after the motor program had run its course (or on the next movement if the response was very short in time, e.g., 200 msec).

Rapid versus slow responses. The present theory extends Adams' (1971) theory dealing with slow, graded, linear positioning responses to movements that are too rapid to allow feedback to be used while the movement is in progress. Theoretically, these two types of responses are controlled and learned in very different ways.

In slow positioning movements, the schema rule determines the expected sensory consequences for the end of the movement specified by the desired outcome. Then the subject moves along the track and compares the response-produced feedback with the expected sensory consequences. He continues moving as long as his errors (the discrepancies between feedback and the expected consequences) are being reduced, and then finally homes in on the target location by reducing his error to zero. Even though the subject is moving actively and some (e.g., Marshall, 1972) have termed these movements "recall," the present position is that the subject is using feedback compared with a *recognition* state, and hence the subject is moving to that location that he *recognizes* as being correct. Such slow responses are dependent on recognition memory and the recognition portion of the schema rule.

What about error detection and the determination of subjective reinforcement after the movement? Clearly, since the subject has used the comparison between expected sensory consequences and response-produced feedback as the basis for determining the endpoint, he will have subjective reinforcement that is zero, reflecting the fact that he moved to that position for which the error signal was zero. He can, therefore, have no further estimate of his own error with respect to the target on that trial unless he receives KR. He cannot, therefore, continue to learn these movements without KR, because he requires KR and/or subjective reinforcement in order to strengthen the response-recognition schema. If the subject lacks this information, the schema remains unstable, and the subject maintains his performance. If the level of practice is low, and the schema is not well developed, he may regress in performance after KR withdrawal because of inappropriate guesses about outcome being paired with the sensory consequences, to the detriment of the accuracy of the schema rule.

Now let us consider rapid responses. Since the time required for exteroceptive and proprioceptive feedback to circle the loop and result in subjective reinforcement is great, rapid movements (such as a task

in which the subject learns to move 30 cm in exactly 150 msec) are completed before the subjective reinforcement can begin to have an effect. Even if it cannot act on that trial, however, the subjective reinforcement is present nevertheless, and still signals the extent to which the expected sensory consequences and response-produced feedback matched, and hence the subject has an estimate of his response correctness after the movement has been completed. This is fed back to the schema as actual outcome, and the rule between outcome and response specifications is strengthened. If KR is withdrawn, the subject has information about response correctness from subjective reinforcement based on proprioceptive and exteroceptive feedback, and performance can be maintained. If sufficient KR practice precedes KR withdrawal, the subject should be able to continue to learn without KR, since subjective reinforcement can provide a substitute for the actual outcome information that KR, when present, provided.

Tasks not requiring accuracy. Many tasks in daily life require one to attempt to do more of something, such as throw harder or jump higher. Theories such as Adams' have difficulty with response evaluation in such tasks, since the perceptual trace is the representation of the "central tendency" of past feedback states. While it is possible that the perceptual trace can provide information about the success of a new jump that is higher than any done previously, accuracy of the subjective reinforcement will be quite low, because the feedback from the new jump is so different from the average of the past attempts. However, in the present theory, the subject can generate expected sensory consequences for such responses on the basis of the recognition schema, and these are not tied to the central tendency of his past consequences. The result is a somewhat better estimate of the extent to which the subject achieved the goal when it lay outside of the range of responses previously executed.

Two Compartments of Memory

While there are two states of memory postulated, they should not be seen as totally

independent, for they are hypothesized to develop using some of the same variables. The recall schema depends on the actual outcome (KR or subjective reinforcement), the initial conditions, and the response specifications, with the rule updated after each trial by an integration of this new information into the existing schema. By contrast, the recognition schema develops using the initial conditions, the sensory consequences, and actual outcome integrated into the existing recognition schema. Thus, while some of the variables are different, both schemata are clearly dependent on the actual outcome and the initial conditions, and both schemata will develop according to the experience the individual has had with these variables. On the other hand, they are separate in that the recall schema uses response specifications and the recognition schema uses obtained sensory consequences. The degradation of feedback when KR is present should have no effect on the recall schema, since KR is used as the estimate of actual outcome, but should have detrimental effects on the recognition schema.

As was pointed out earlier, Adams' (1971) theory has difficulty in explaining how one could continue to learn a positioning task after KR had been withdrawn, or even how one could simply maintain performance without KR. The problem was that if the subject ever made an error, the perceptual trace would be degraded somewhat, he would consequently make a greater error on the next trial, the trace would be further degraded, and so on until movement had become very inaccurate. For the slow movements, the schema notion cannot predict continued learning without KR since subjective reinforcement always indicates that the subject has moved to the correct location on each trial. The present theory can, however, predict that performance in slow tasks should be maintained, since in order to change the recognition schema the subject requires the pairing of actual outcome and the sensory consequences. Of the two he has only the sensory consequences after the movement, and hence no changes in the schema can occur. In the cases where KR is withdrawn after very few KR trials, dec-

rements in performance might occur because of incorrect guesses about the actual outcome, with inappropriate updating of the schema resulting in the performance decrements commonly seen (e.g., Bilodeau, Bilodeau, & Schumsky, 1959).

In rapid movements, however, the KR-withdrawal effects are somewhat different. When enough practice with KR has been provided so that the subject has an accurate recognition schema, if KR is withdrawn the subject has the capability to store all that is necessary for continued recall schema development: the initial conditions, actual outcome (from subjective reinforcement), and response specifications. Hence, if the subject makes an error, it does not degrade the recall schema mechanism, but rather strengthens it, since in an error response actual outcome and response specifications will be lawfully related in the same way that they are in more nearly correct movements. The result is that the recall schema is at least maintained in strength, and is perhaps increased in strength, with KR withdrawal. Of course, the recognition schema remains stable because it requires the pairing of actual outcome and sensory consequences; in this case the actual outcome and sensory consequences are redundant, since the former was computed on the basis of the latter, and no updating can occur.

SOME SUPPORTING EVIDENCE

This section provides some experimental evidence that is taken as support for the various hypothetical processes assumed to occur in the performance and learning of discrete tasks. In addition, independent variables that are hypothesized to influence the various processes are mentioned, and experimental paradigms for subsequent tests of these ideas are suggested. Before turning to the evidence, however, a discussion is necessary concerning the dependent variables assumed to be indicative of the processes and states proposed.

Dependent Variables

Motor recall. Recently there has been considerable debate about which of many error scores should be used as measures of

performance in tasks demanding accuracy. While the issue is certainly not new (e.g., Woodworth, 1938), Schutz and Roy (1973) have reopened the argument, advocating the use of constant error and variable error as measures of accuracy, arguing that absolute error should not be used because it is confounded with constant error and variable error. It is appealing to have measures that are independent within subjects such as constant error and variable error, allowing variable error to represent consistency and constant error to represent directional biases, but it can be argued that absolute error is to be preferred as a measure of performance. First, having two dependent measures in one experiment can allow opposite conclusions to be drawn about the processes that they are thought to estimate (e.g., Dobbins & Rarick, 1975), leading to equivocal findings. Second, absolute error has a solid history of use in psychology and motor behavior, probably because it is intuitively meaningful; it is the amount by which the subject was incorrect in his movement. Third, one can imagine the performance of an individual who is told that he will be financially rewarded for reducing his constant errors on a series of trials; he will probably move on trial n to cancel out a constant error made on trial $n - 1$, and will not attempt to achieve a correct movement. It seems that individuals try to reduce the amount by which they are incorrect in the task (i.e., the absolute error), and measuring constant error and variable error when the subjects are not aware of it could provide misleading data. For these reasons, the use of absolute error as a measure of recall-schema strength seems justifiable.

In some of the experiments cited in this article, variable error is used as a substitute for absolute error, and it can be argued from Schutz and Roy's (1973) analysis that the variable error and absolute error are probably measures of the same state or process, since the constant errors in such tasks are nearly zero. Finally, Henry (1974) advocates the use of the total variability of the subject's scores around the correct value as a measure of overall accuracy in responding, and this measure is probably highly associ-

ated with both absolute error and variable error under conditions where constant error is nearly zero.

Motor recognition. A number of researchers (e.g., Newell, 1974; Newell & Chew, 1974) have argued that the appropriate statistic for the estimation of recognition memory strength is the absolute value of the difference between the objective and subjective error, and that the within-subject correlation does not indicate the extent to which the subject is incorrect in his judgments, only that he is sensitive to the direction of his errors. They argue that he could be in error by a constant of 100 msec (i.e., his average guess might be 100 msec too large) and yet this fact would not be represented by a drop in the correlation. While they are correct on this point, there are a number of reasons that the correlation is to be preferred to the difference score.

First, there has been negative reaction to difference scores (e.g., Bereiter, 1963; Schmidt, 1972a), and the objective-subjective difference score is not free from undesirable effects, which include (a) an extreme correlation with the minuend, which in this case is the absolute error ($r = .80$ in Schmidt & White, 1972), (b) probable nonlinearity of relationship with the construct (in this case, the recognition-schema strength) that the difference score attempts to measure, and (c) statistical unreliability due to the summation of errors in measurement for the two components. A further difficulty for the present theory is the fact that the objective-subjective difference and the absolute error theoretically are measures of separate states of memory; however, the method of computation of these two statistics allows the inclusion of absolute error in both measures, leading to a strong correlation between these two scores, even to the point that the plots of these two scores over trials yields nearly identically shaped functions. In such cases, it would be possible to show that a given experimental variable influences both the recognition and recall measure identically because of a statistical artifact, leading to the erroneous conclusion that recall and recognition were not separate.

The objective-subjective correlation is to be preferred, because it is clearly sensitive to the direction of the subject's guesses. It also counters a strategy employed by some subjects whereby they guess the correct score (e.g., 200) on each trial of a 200-msec task. In such cases the correlation is zero, reflecting a total insensitivity to the size and direction of the error, whereas the difference score is minimized (especially if the constant error is zero for that subject), making him appear to be quite talented in predicting his score. As a measure of the strength of the recognition schema, therefore, the correlation statistic is to be preferred, but could be supplemented with the average objective-subjective difference to be certain that there were not large constant errors in guessing.

Support for the Schema Notion

The most important portion of the present theory is the schema, and this section deals with experimental and other evidence for it. The recall and recognition schemata will be considered separately.

Recall schema. Strictly speaking, very little published work supporting a schema for response production can be found. However, there are a number of studies suggestive of the schema's existence, and there is subjective-anecdotal evidence as well.

Anecdotal evidence for the schema is that subjects can produce sequences of movements that they have never performed before. A classic example concerns handwriting, in which a person's signature can be recognized as his regardless of the size of the actual marks produced; one can produce the same signature on a check or 10 times larger on a blackboard (Merton, 1972, p. 4). Of course, quite different musculature is used in these two signatures, with the movements confined to the fingers and perhaps the hand in small writing, and the entire arm in larger writing. The schema notion explains this by saying that the movements are all run off by a large motor program that needs to have certain specifications in order to produce a given movement sequence. When the specifications are that the movement should be small, rapid, and slanted

slightly, the program can carry out these movements as planned. Of course, a motor program theory such as that of Henry (1960) cannot handle these findings, since it postulates that the movements are the result of a specific motor program, with a different program needed for each movement "style." The fact that individuals can do this seems to be evidence directly in conflict with such one-to-one program-movement-style ideas.

Pew (1974) adheres to the idea that the movement is programmed rather generally, and then the movement parameters are specified before the movement is initiated. In a study by Armstrong (Note 3) a sequence of lever movements was learned in a number of sessions. An important result was that the timing in the sequence appeared to be generalized, in that when the movement was done too rapidly the entire sequence was speeded up, keeping the temporal relationships among the various submovements approximately constant. The implication that Pew (1974) draws from these data is that timing is one of the response specifications that serves as a "parameter" of the motor program.² Also, Glencross (1973) found in handwheel cranking that the movements were very similar in terms of the timing of the onsets of force application even when resistance was added and the radius was changed; he labeled the phenomenon "gradation of effort." It is as if there was a program for cranking, and that changing the quality of cranking (i.e., less force, more speed, etc.) provided a situation in which the subject still used the program, but with a different set of specifications. While the present article does not attempt to deal with such continuous responses, these findings at least suggest that a schema-like process could exist in discrete movements as well.

However, the most impressive kind of evidence that could be generated in support of the schema is that subjects can produce movements of a given class that they had, strictly speaking, never performed previ-

ously; an example is the basketball player who shoots from various places on the floor with great accuracy. The notion is that the varied previous shooting experiences led to increased schema strength, providing a basis for generating novel movements of that same class. One important prediction is that increasing either the amount or the variability of such previous experiences lead to increased schema strength. These predictions suggest a test of the schema notion in terms of transfer of learning.

Experiments testing this idea have been conducted, and some of them are quite old (e.g., Crafts, cited in Ellis, 1965; Duncan, 1958). For example, Duncan used a task in which the subject had to position a lever into one of 13 slots depending upon which of 13 lights was illuminated, with the movements being done as quickly as possible. The task could be varied to produce 12 similar versions, and subjects received either 1, 2, 5, or 10 variations in a training session, with the total number of trials held constant. Duncan found that the amount of transfer to two novel variations of the task was a positive function of both the amount and variability of training, especially as the variability was increased from one to two tasks. While this study supports the predictions from the theory strongly, it could be argued that what was learned and transferred were cognitive-conceptual relationships related to which light went with which slot, and the findings may not tell a great deal about the existence of schemata for producing movement. To support the recall schema idea, a task should be used that involves primarily motor learning, with minimized cognitive aspects, so that transfer as a function of response variability could be attributed to the development of schemata that define the response specifications.

Schmidt and Shapiro (unpublished study, 1974) used Duncan's design with a task that could be considered more motor in nature than his light-slot task. The task involved moving the preferred arm and hand to knock over four small barriers in a predefined order, and minimized movement time was the subject's goal. The task was varied by

² See also Brooks (1974), who found that timing appeared to be a parameter of alternating hand movements in monkeys.

changing the locations (but not the orders) of the barriers, so that four tasks with slightly different movement-segment lengths and intersegment angles were defined. One group performed three tasks (40 trials of each with KR) while a second group performed one task (120 trials with KR). When subjects were transferred to a fourth novel version of the task, there was some tendency for the group with high variability (three tasks) to perform more quickly than the group with low variability (one task), with these differences increasing somewhat over subsequent practice with KR, but the differences were small and did not reach statistical significance.

The failure of the Schmidt and Shapiro study to support the schema notion does not, of course, disprove it, as there are a number of alternative explanations. First, the task involved the preferred arm and hand with college-age subjects, and it is possible that schemata for movement production had already been developed throughout extensive previous movement experience with that limb. This suggests that future work along these lines should use limbs and movements that have not been used extensively, or should use younger subjects in whom the schemata have not been developed. Also, it is possible that the variations of the task did not produce movements that were sufficiently different, or produced movements that were different in the "wrong" ways.

Recognition schema. As was mentioned in a previous section, there is considerable evidence in the dot-pattern recognition studies for the existence of schemata. This evidence indicates that subjects not only can store the individual patterns seen, but also can store some abstraction of the patterns. This abstraction is the schema and consists of a set of rules for determining whether a given new pattern is or is not a member of the class of objects described by the schema rule. As strong as this evidence seems to be, there are two essential ways in which the experimental situation dealing with dot patterns differs from the motor recognition schema in the present theory.

First, the schema in the dot-pattern work represents the central tendency of the past

sets of stimuli seen. This is different from the present situation in that the recognition memory for, say, a novel ball throw may not represent the central tendency of all balls previously thrown. To have it represent the central tendency would provide a recognition state that would not be associated with a new response, that might not be near the central movement pattern, and that would not provide recognition for novel movements. Rather, what the schema represents is the relationship among a number of sources of information stored on each trial, and in the motor recognition case, these are the actual outcome and the actual sensory consequences. Hence, the schema notion is extended from becoming the central tendency of a set of stimuli to the relationship between paired members of two sets of stimuli.

A second and less cautious extension of the recognition schema notion is to types of stimuli other than the visual patterns studied by Posner and Keele (1968, 1970) and others, and it is assumed that such abstracting processes can occur with auditory, tactile, and kinesthetic stimuli as well. There is some support for recognition schemata with kinesthetic stimuli in a report by Williams and Rodney (in press). As mentioned earlier, blindfolded subjects attempted to learn a criterion position along a trackway in two different ways. Group TO (target only) moved 16 times to a stop at the criterion position, whereas Group IR (interpolated random) moved to 16 randomly ordered stops that surrounded the criterion position; Group IR was instructed that the criterion position was in the center of the 16 positions presented. The critical transfer test (see Figure 1) showed Groups TO and IR performing with nearly equal absolute errors on the first block, but Group IR maintained performance over subsequent blocks, whereas Group TO regressed sharply in its performance.

The interpretation in terms of the schema theory is that the 16 practice positions performed by Group IR enabled subjects to develop a recognition schema for the locations along the trackway; the reduced variability for Group TO did not permit this schema

development. Then, on the transfer trials, Group IR subjects could generate the expected sensory consequences associated with the criterion position (even though they had not experienced this position previously), and could then move so as to match the incoming proprioceptive feedback with the expected proprioceptive feedback. The recognition-schema strength was approximately maintained over the five blocks for Group IR, but for Group TO the traces associated with the criterion position apparently faded, resulting in less accurate responding for these subjects. The important point is that Group IR subjects could recognize the criterion position without having experienced it previously, and they could recognize it as effectively as Group TO subjects with experience at the criterion location. While the experiment does provide support for the schema theory, it is perhaps more important because it provides strong contrary evidence to the prediction from Adams' (1971) theory, which says that Group TO with practice at the criterion position should have greater accuracy (due to increased perceptual trace strength) and should maintain accuracy longer than Group IR. These findings provide considerable difficulty for Adams' position.

Separate States of Memory

There is considerable evidence in the verbal learning literature to indicate that there are two separate states of memory, one for recall in which subjects must produce the response term, and one for recognition in which subjects must only say whether the item has been seen before. (See Underwood, 1972, for a review of this issue.) The same view is presented in this article and is in general agreement with Adams (1971). Under this view, recall memory is the state that produces the movement, and recognition memory is the state that determines whether the subject can recognize that the movement was correct. While there is strong evidence for this view in verbal learning, what is the evidence for this contention in motor skills?

A research strategy. It was realized long ago in verbal behavior that recognition measures (in terms of number correct) were

usually larger than were recall measures, but that this difference, in and of itself, did not necessarily mean that the processes underlying recognition and recall were different. If it could be shown that a given experimental variable influenced recognition in one way (e.g., increased it) and influenced recall in another way (e.g., decreased it), this would provide evidence that the two memory states follow different laws. If two systems have different laws, it makes little sense to maintain that they are the same system, and finding that some experimental variables (e.g., intentional versus incidental instructions) affect recall but not recognition has been used as support that recall and recognition are separate systems.

Motor recognition and recall. The recognition-recall issue had not been studied in motor skills until Adams (1971) proposed the independence of these states in his theory. Since the introduction of this theory, however, there have been a few attempts to show that certain variables influence recall and recognition differently in motor skills. One such attempt was made by Schmidt et al. (in press), who sought evidence for the independence of recall and recognition using KR delay as the main independent variable. In Experiment I, subjects learned to recognize lifted weights, a performance reasoned to be based on recognition memory. Increased KR delay (25 sec vs. 5 sec) caused a drop in the correlation between the actual weight and the subject's guess from .75 to .46, supporting the conclusion that KR delay is a variable in recognition memory. In Experiment II, the authors used a rapid ballistic task (as with Schmidt & White, 1972), attempting to show that KR delay was a factor in the recognition memory for the task (as it was with Experiment I) but was not a factor in the recall memory.

The task in Experiment II involved a paradigm wherein the subject was asked to learn to move a horizontal slide 30 cm (with follow-through permitted) in 200 msec. The subject made a response, the experimenter asked for his guess about his error (subjective error), and then the experimenter provided KR in terms of the actual (objective) error. The experimental variable was again

TABLE 1
RECALL AND RECOGNITION MEASURES AS A
FUNCTION OF DELAY OF KNOWLEDGE OF
RESULTS IN ACQUISITION

Measure	Delay interval	
	5 sec	25 sec
Variable error	18.6 msec	20.4 msec
Constant error ^a	+14.4 msec	+16.3 msec
Objective-subjective correlation	.92	.61
N	20	20

Note. From "Some Evidence for the Independence of Recall and Recognition in Motor Behavior" by R. A. Schmidt, R. Christenson, and P. Rogers, *Journal of Motor Behavior*, in press. Copyright 1975 by Journal Publishing Affiliates. Reprinted by permission.

^a Positive values are overshoots.

the delay of KR. The rationale was that if KR is delayed and the proprioceptive feedback received from the movement fades with time, the actual consequences (based on KR) cannot be paired effectively with the faded sensory consequences, and hence the recognition schema will not develop efficiently. Recall, or the production of the movement, is dependent upon KR and the response specifications (which, it is assumed, do not fade with time as rapidly as the sensory consequences), and hence the theory would predict that the measure of response production would not be influenced by the delay of KR.

Subjects were given 50 trials, with KR provided either immediately (after 5 sec) or delayed for 25 sec, and then practiced 30 additional trials with no KR. The measure of recall in this task was the accuracy in terms of the 200-msec goal, and either constant error or variable error was the recall indicant here. Recognition concerns the extent to which the subject can detect his own errors, and the objective-subjective correlation was used as the indicant of recognition memory. The test of the strength of the recall memories was carried out on the no-KR trials in which the level of the independent variable (KR delay) was equated across groups, in keeping with the established paradigms for determining the learning versus performance status of various experimental variables (e.g., Schmidt, 1975).

The experiment supported the independence of recall and recognition when it was

found that increasing KR delay had no significant effect on recall measures (constant error or variable error), but had significant detrimental effect on the recognition indicant, the objective-subjective correlation. As can be seen in Table 1, the 5-sec delay group displayed correlations of .92, whereas the 25-sec delay group showed a correlation of only .61, representing a significant drop in the error-detecting sensitivity of the long-delay group. Therefore, KR decay appears to be a variable that causes decrements in recognition, but leaves recall unaffected, and these findings are in the direction predicted by the theory.

If it is true that KR delay is a variable in the development of the recognition schema, then the learning of any task dependent on the recognition schema should be retarded as well. One such task is linear positioning, which, according to the theory, is dependent on the match between the expected sensory consequences (generated from the recognition schema) and the incoming proprioceptive feedback during the movement. The difficulty with all this is that there is ample evidence that linear positioning tasks are not affected by KR delay (see Bilodeau, 1966, for a review). One possibility is that the evidence on linear positioning has used procedures that prevented the matching of feedback and the expected sensory consequences because of the time constraints usually placed upon subjects, and one could assume that subjects may have switched to a strategy wherein they program the movements. In this case, the production of the movement would be dependent upon recall memory and the recall schema. Schmidt and Shea (unpublished study, 1974) tested this idea by insisting that subjects move very slowly and carefully, encouraging the subjects to "hunt around" for the correct location and preventing programmed movements by varying the starting location; but KR delay still had no effect on learning in this task. Newell³ has pointed out that it is possible that KR delay does not affect the recognition schema, but

³ The author wishes to thank Karl M. Newell for providing this suggestion in a personal written communication, October 1974.

rather affects the development of the error-labeling schema; if the subject could not label the errors effectively in the Schmidt et al. (in press) study, we may have incorrectly come to the conclusion that the recognition schema was impaired. More work will be needed to resolve this question.

Even if the Schmidt et al. (in press) study may be inconclusive concerning the recognition-recall distinction, there are other data that provide somewhat stronger evidence. The same basic strategy was used by Newell and Chew (1974) and Schmidt and Wrisberg (1973), but here the independent variable was the amount of response-produced feedback. In both studies, the presence of combined vision and audition was varied in the rapid, ballistic task (Schmidt & White, 1972), and they showed that vision and audition had no effect on response production in terms of absolute error. But in Schmidt and Wrisberg (1973), when KR was withdrawn, subjects with vision and audition regressed far less in performance than did subjects without these feedback sources present; the interpretation was that vision and audition enabled subjects to generate accurate subjective reinforcement via the comparison of the feedback with the expected sensory consequences from the recognition schema. Also, Newell and Chew (1974) showed that eliminating vision and audition in a transfer test increased the difference between the objective and subjective errors, a measure taken to be sensitive to the strength of the recognition schema.

The conclusion from this section is that variables have been isolated—the quality of feedback and perhaps KR delay—that influence the recognition memory for motor tasks but do not affect measures of the recall memory. This, plus the evidence from the verbal learning literature, provides support for the hypothesized separation of recall and recognition processes in motor skills.

Increased Error Detection With Practice

One of the main predictions of the theory is that error detection will increase with KR practice, and Schmidt and White (1972) and others since then (Newell, 1974; Newell & Chew, 1974; Schmidt & Wrisberg, 1973)

have shown this quite clearly. In the Schmidt and White experiment, subjects were given 170 trials of the ballistic slide task, with a target movement time of 150 msec. Subjects provided subjective errors prior to the experimenter giving them their objective errors (i.e., KR). The correlation increased steadily over the course of practice, as can be seen in Figure 4, with the correlations being quite low (around .20) in initial practice and rising to around .90 in later practice. These data show clearly that subjects increased their sensitivity to the direction of their errors with practice, and support the notion that the schema for response recognition was increasing in strength.

Error Labeling System

What is the evidence for such a labeling system? Subjectively, it seems rather obvious that such a system must exist, since subjects clearly transfer information from the neurologically coded sensory input to words about the error, and they can do this rather well. Schmidt and White (1972) found objective-subjective correlations of .90, and subjects were able to guess their errors within 7 msec of the actual time moved on the average. Thus it seems clear that subjects can estimate their own errors quite well.

However, additional data arguing for a separate status of the two error signals (sensory errors versus subjective reinforcement) were presented by Newell and Boucher (1974). They asked blindfolded subjects to move a linear slide 20 times to a stop, then to estimate the movement distance verbally, and then to move to the designated position once without the aid of the stop. One group estimated the distance in millimeters and another group in inches, with all subjects having had more experience with the inches scale. The guesses in inches (transformed to millimeters) displayed nearly half the error of the guesses in millimeters, while the accuracy in moving to the target was not significantly different between groups. The interpretation is that the sensory signal arising from the comparison of actual and expected feedback was identical for the two

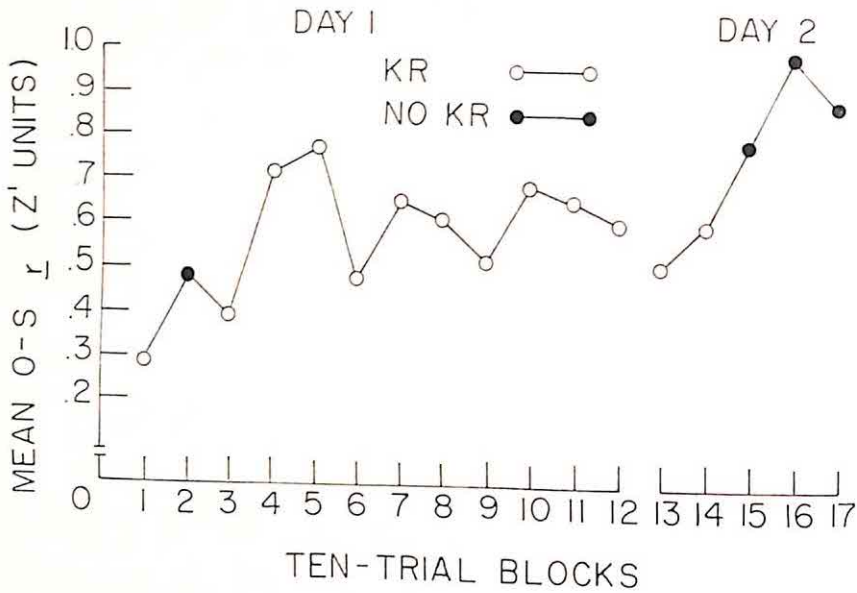


FIGURE 4. Mean objective-subjective (O-S) correlation as a function of practice. Abbreviation: KR=knowledge of results. (From "Evidence for an Error Detection Mechanism in Motor Skills: A Test of Adams' Closed-Loop Theory" by R. A. Schmidt and J. L. White, *Journal of Motor Behavior*, 1972, 4, 143-153. Copyright 1972 by Journal Publishing Affiliates. Reprinted by permission.)

groups, but that the more efficient schema for labeling in terms of the inches scale resulted in a more accurate report of the distance. These data, then, provide some support for the idea that there exist two states of error: raw sensory signals and labeled subjective reinforcement.

Learning Without KR

The predictions of the present theory for no-KR learning differ considerably for rapid tasks (i.e., with movement time less than 200 msec) and slow positioning tasks. The categories of movement are discussed in turn.

Positioning tasks. The theory hypothesizes that the subject develops a recognition schema over the course of practice with KR, and that he uses the comparison of actual feedback and the expected sensory consequences generated on each trial as the mechanism by which he arrives at the correct location. The evidence is quite clear that without postresponse error information in the form of KR (and without presenting the correct movement on each trial), no learning can occur without KR; this conclusion is

supported by a number of investigations (e.g., Bilodeau et al., 1959), and the evidence has been summarized by Bilodeau (1966). This is entirely consistent with the present theory, since the sensory consequences of the correct response are required to strengthen the recognition schema, and the subject has no way of learning the correct location without the guidance provided by KR.

Adams (1971) suggested that subjects could continue to learn these responses without KR after sufficient KR practice was provided, during which the perceptual trace (the expected sensory consequences in the present terms) was strengthened. The present theory does not predict this, and there are no data to suggest that such should be the case in linear positioning tasks. Theoretically, after the recognition schema is developed, withdrawing KR provides a basis for continuing to respond with the former level of accuracy, but any improvement, except by accident, is not possible. Subjects use the recognition memory to produce the movement, attempting to match their current feedback with the expected sensory conse-

quences, but as discussed earlier, their subjective reinforcement indicates that they had moved to that position for which they received a match and for which the subjective reinforcement was necessarily zero. Hence, their subjective reinforcement is not capable of signaling their actual movement distance on that trial, and the substitution of subjective reinforcement for the withdrawn KR is ineffective. Hence, after the movement subjects have the actual sensory consequences to store for use in the updating of the recognition schema, but they do not have the other essential ingredient—the actual outcome. Therefore, no further development of the schema can occur, and learning does not continue after KR withdrawal. These predictions are upheld by a number of separate investigations reviewed by Adams (1971).

The reader should take cautiously the conclusion that if postresponse error information in the form of KR is not available in such tasks, learning cannot occur, for there are two studies that show clearly that learning *can* occur without KR under certain conditions. Solley (1956) used a motor-driven tilt-chair that the subject could operate by a two-way finger switch. The subject was firmly strapped to the chair and the head was immobilized with a bite-bar. He was powered to the vertical position, whereupon the experimenter said "vertical." Then the experimenter tipped the subject 30° to the left or right, and the subject attempted to replace himself in the vertical position with appropriate switch movements. When the subject had achieved what he thought was vertical, Solley "randomly" moved him about to disguise the relationship between vertical and the position to which he had actually moved, and then finally moved him to vertical again and again said "vertical." The subject was then tipped 30° and attempted to again replace himself in the vertical. Over the course of 30 such trials, the absolute errors in arriving at vertical decreased markedly and significantly, and this was achieved without ever providing direct error information. The crucial difference between this study and those referred to earlier is that the correct location was presented on every trial, and apparently this is

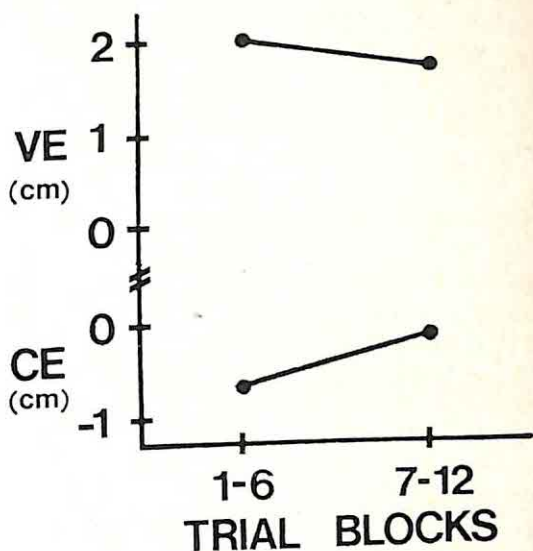


FIGURE 5. Mean constant and variable errors (CE and VE) as a function of practice. (From "A Note on Improved Motor Performance Without Knowledge of Results" by C. A. Wrisberg and R. A. Schmidt, *Journal of Motor Behavior*, in press. Copyright 1975 by Journal Publishing Affiliates. Reprinted by permission.)

all that is necessary to acquire the correct movement.

Wrisberg and Schmidt (in press) duplicated the Solley study with a linear positioning task. They presented the correct location on each trial by having the subject move to a stop, and then waited either 5 or 50 sec before having him attempt to move to this position without the stop. Subjects repeated this trial sequence of the correct movement plus an estimation 12 times to the same target, and KR was never provided as to the response correctness. The results of the study are shown in Figure 5, where the constant errors and variable errors are presented for two successive 6-trial blocks. Both constant error and absolute error (not graphed) decreased significantly with practice, with less undershooting as a function of trials; variable error decreased as well, but the differences just failed significance. These two studies provide evidence that subjects can learn slow positioning tasks without postresponse error information, provided that the correct location is provided on each trial. Adams and Dijkstra (1966) produced similar findings in linear positioning.

The present theory can handle these findings quite easily. When the subject is presented with the correct movement, he stores the sensory consequences along with the desired outcome, and the recognition schema begins to be developed. Then when he attempts to reproduce the movement without the stop, he moves to that position he recognizes as correct. On the next trial he receives the sensory consequences of the correct response again, further updates the schema, and reproduces the second trial with greater accuracy than the first. KR is not necessary to learn because subjects received the appropriately paired sensory consequences and desired outcome (the latter being a constant) on each trial, and this is all that is needed for development of the recognition schema.

Rapid movements. The predictions about learning without KR in rapid tasks are somewhat different from those in positioning tasks. Theoretically, the movement sequences are carried out totally open loop in the rapid tasks, with recall memory having a major role in movement production. Recognition memory and the resulting subjective reinforcement occur after the movement, since these feedback loops are considered quite slow. Therefore, in such tasks there is the capability of detecting one's errors after the movement so that corrections can occur on the subsequent movement.

Even so, however, considerable initial practice is needed in order to strengthen the recognition schema to the point that it can provide subjective reinforcement of sufficient accuracy to be of advantage to subjects in the absence of KR; hence, the theory cannot predict learning "from scratch" without KR, except perhaps in the situations where (a) the correct movement is presented in a way analogous to the method used in the Solley (1956) and Wrisberg and Schmidt (in press) studies, or (b) the task to be learned is a member of a highly practiced class of motor responses for which a strong schema has been previously developed. There is apparently no evidence concerning the former situation, but the latter possibility has some support from Henderson (1974), who showed (with experienced darts players)

improved performance in dart throwing without any KR or other feedback information.

After sufficient practice with KR, the subject not only should be able to continue to perform effectively under KR-withdrawal conditions, but also should be able to continue to learn the movement. This prediction, however, has been quite difficult to support because of a number of methodological problems. In order that the task have strong recognition memory, it should presumably be simple enough that the recognition memory can be developed within the time constraints of the typical experiment. However, with such simple tasks, the improvement in performance is quite rapid, with nearly asymptotic conditions reached in very few trials (e.g., Schmidt & White, 1972; Schmidt & Wrisberg, 1973). If KR is now withdrawn, subjects have no room for improvement, because errors are already at an effective floor in the task. Thus, attempts to demonstrate continued learning without KR have been almost universally unsuccessful. An example is the Schmidt and White (1972) study in which KR was withdrawn after 150 trials in a 150-msec movement-time task; while there was some indication of a continuing downward trend in absolute errors, the effect was very small and was not statistically reliable.

However, an alternative approach has proven useful. Schmidt and McCracken (unpublished study, 1974) allowed subjects to perform 200 trials of rapid movement with a 200-msec movement-time criterion with KR on one day, and then brought subjects back 1 week later, providing 20 additional trials with no KR at any time. Through the forgetting that would occur over the 1-week interval, subjects could return to the task with elevated error, allowing some room for subsequent decreases, and thus the study investigated relearning without KR. The 20 trials on the no-KR period were grouped into blocks of 5 trials and are presented in Figure 6. The decrease in absolute error was significant, but there were no significant changes in either constant error or variable error. While these downward trends were significant, the error at the

end of this period was considerably larger than it was at the end of initial KR practice, and whatever relearning took place was quite slow and rather incomplete. Nevertheless, these data do suggest that relearning can occur without KR when initial KR practice has been provided (although, strictly, a warm-up interpretation is possible), and they support the theoretical expectations about no-KR learning.

The interpretation is that the recognition and recall schemata were retained to some extent over the 1-week layoff, but that they both suffered considerable weakening. The error on the initial no-KR trial was quite large, suggesting that the choice of the response specification was degraded by forgetting. Even though the schemata had undergone forgetting, it appeared that sufficient subjective reinforcement was available to allow improvement over the next 19 trials. Further work needs to be done using more extensive practice with KR, and with more no-KR trials to determine if this prediction might hold more strongly if given the chance.

Response-Produced Feedback and Recognition

The present theory hypothesizes that the development of the recognition schema, and hence the capacity to detect errors either during (slow responses) or after (rapid responses) the movement, is dependent upon the pairing of the actual outcome (based on KR) with the sensory consequences of the movement. Attempts to manipulate proprioceptive feedback in motor tasks have been fraught with difficulty arising from either unknown effects of direct degradation of feedback, as in the "cuff technique" (Laszlo & Bairstow, 1971), or by the confounding effects of changing the dynamics of controls (e.g., Bahrck, 1957). If the manipulation of feedback can be argued to change the movement required of the subject, then one cannot be sure that degraded error detection after such treatments was due to the effects on the recognition schema or due to the fact that subjects had been attempting to learn a different task. Given these limitations,

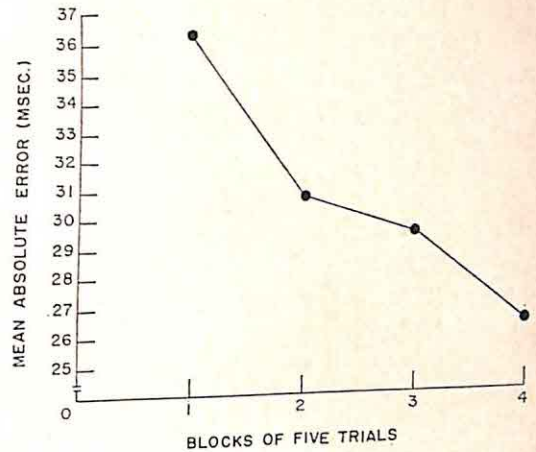


FIGURE 6. Mean absolute error in relearning without knowledge of results (from Schmidt & McCracken, unpublished study, 1974).

however, there are a few studies that link feedback and error detection.

Within the framework of Adams' (1971) theory, a number of authors attempted to show that error detection and error correction were enhanced by the enrichment of feedback and practice. Adams & Goetz (1973) showed that providing vision, audition, and increased force cues in linear positioning (as compared with no vision and audition and minimized force cues) increased accuracy in detecting and correcting errors. Practice was effective in increasing the recognition processes as well. Adams, Goetz, and Marshall (1972) have provided similar findings. The interpretation is that the recognition schema uses the proprioceptive information to form the rule, and that enhancing this information renders the rule more effective.

With more rapid movements, Schmidt & Wrisberg (1973) showed that eliminating vision and audition had no effect on performance when KR was present, but that when KR was withdrawn after 20 KR trials, the groups without vision and audition progressed in their performance to a greater extent than groups with these information channels present. The implication was that KR practice when vision and audition were present permitted the establishment of a recognition schema based on these channels, and that no-KR practice then allowed subjects the opportunity of using subjective re-

inforcement based on vision and audition; groups without vision and audition in KR practice regressed more because of less accurate subjective reinforcement. An interesting finding was that while earlier experience with vision and audition had strong effects on KR withdrawal performance after 20 trials, there were no such effects after 75 trials, suggesting that the channels of information used by the recognition schema may change somewhat as practice continues; vision and audition may have been important in earlier error-detection performance, and perhaps proprioception had become more important later. There is certainly more room for experimentation on this question.

ROLE OF EFFERENCE COPY

An important line of thinking in motor skills research today is the notion of efference copy. There are a number of points of view about how it might operate and the issues are definitely not settled. The role of efference in the present theory, therefore, is somewhat unsettled also, and this section attempts to review some of the important issues and to explain the thinking that led to the present position taken on efference. Space does not permit a thorough treatment of this area, but for reviews consult Gyr (1972) or Jones (1974).

Evidence for Efference

The notion of efference copy began with the eye-movement work of Helmholtz (1925) and later Holst (1954). These writers argued that for accurate visual perception it was essential that the central nervous system have information about motor commands sent to the muscles that move the eye. Without this information, the visual images are ambiguous, in that the organism cannot know whether a shifting visual image means that the eye was moved in a stable environment or whether the eye was stationary in a moving environment. This finding led Holst (1954) to propose that an image of the motor commands (an efferent copy or corollary discharge) is "fed forward" to modify the incoming visual sensations to compensate for the fact that the eye was ordered to move. This contention has been strengthened

recently by neurological evidence that such pathways exist (e.g., Chang, 1955; Li, 1958).

From this basic idea for visual perception, a number of other versions of the idea were created, and the application was made to the perception and production of movement. One major problem with the idea of efference for movement control is that the same term has been used to represent a number of very different ideas. There are at least three ways that the term has been used in the literature.

Efference as a feed-forward process. The most general, and perhaps most widely accepted, version of efference copy is that it consists of information that is fed forward before or during a movement that prepares the system in some way for the receipt of various response-produced feedback signals. The feeding forward of gamma efferent information to the spindles so that the movement is carried out as ordered, and the information fed forward so that accurate visual perception can occur are examples of this kind of idea. Generally speaking, such processes are thought to occur at a number of levels within the central nervous system. There is no necessity that the information fed forward be a literal copy of the motor commands sent to the musculature, as with the following two models.

The Holst idea. In the original statement of this model, a literal copy of the motor commands sent to the eye muscles is fed forward to modify the incoming visual information from the retina (Holst, 1954), enabling the accurate interpretation of the visual signals. The idea has been extended to movement control by assuming that an efferent copy of movement commands is compared against the inflow of proprioceptive feedback, with resulting mismatches representing an error in the movement. Jones (1974) has termed this idea the "inflow model" because it depends upon the inflow of proprioceptive feedback.

One problem with this inflow model is that the proprioceptive feedback and efference copy are in different "languages," with proprioceptive feedback coded in terms of the movements of joints, and efference coded in terms of the commands to the musculature.

ture. With the information in different codes, how, strictly speaking, could these two states ever match? A solution often heard is that massive recoding must exist throughout the nervous system, but this begs the question, as it does not provide testable assertions about how such recoding is accomplished.

A second problem, though, concerns the fact that such a mechanism cannot provide information to the subject about the extent to which he has achieved the environmental goal for the movement; it can only signal the extent to which the chosen movement was carried out as planned. As discussed earlier, this type of model cannot account for the evidence dealing with the detection of errors in movement (e.g., Schmidt & White, 1972) or the relearning without KR (Schmidt & McCracken, unpublished study, 1974).

The Jones idea. Largely because the inflow model depends on proprioception, Jones (1974) rejected this notion and argued in favor of what he terms the "outflow model" where a literal copy of the efferent commands is monitored centrally without comparison with proprioceptive feedback. According to the argument, knowing what the limbs were commanded to do (and knowing that the limbs will carry out these commands) is sufficient information to be able to determine where the limbs are after the movement. Hence, this version of efference has been used to explain the perception of limb locations in space. This basic idea, however, came from the eye-movement work where the loads on the eye musculature were nearly constant, permitting one to have information about the eye's location from knowledge of the efferent signals (e.g., Festinger & Canon, 1965). But the loads on the skeletal musculature are not nearly as predictable as they are in the eye, and it is difficult to imagine how limb location could be sensed solely from the record of the efferent commands.

In spite of the argument concerning the unpredictability of limb loading, some recent evidence has been interpreted in terms of an outflow model, because of the apparent difficulty of feedback models to account for the

findings. These concern (a) the learning of movements in deafferented animals (e.g., Taub & Berman, 1968), and (b) the correction of movement errors in less than one reaction time (e.g., Angel, Garland, & Fischler, 1971).

Taub and Berman (1968) used monkeys that were deafferented from the head down in a shock-avoidance experiment. They prevented the animals from seeing their limbs by the use of a large collar extending outward from the neck, secured them into a chair-like structure, and taped a pneumatic bulb into the palm of one hand. When a shock was administered to a normally afferented portion of the head, the animals learned to turn off the shock by squeezing the bulb. The fact that the animals could learn a hand movement without any feedback from the periphery was surprising to many theorists, and the leading interpretation has been in terms of the outflow notion. According to the argument, (a) feedback is essential to learning, (b) there was no sensory feedback, (c) outflow or efference copy could provide information about the commands issued, (d) therefore, outflow was the mechanism that mediated learning.

The second line of evidence has used two-choice step-tracking tasks, in which the subject moved in one of two directions to a target as quickly as possible after a stimulus. Occasionally, the subject would make an error, in that he would initially move in the improper direction. When this happened, he was to reverse his direction as quickly as possible, and move back in the correct direction. Angel and his colleagues (Angel et al., 1971; Angel & Higgins, 1969; Higgins & Angel, 1970) and Megaw (1972) found that the latency of these corrections (the time from the first movement in the wrong direction until the subject began to correct his movement) was sometimes very short (from 60 to 90 msec), clearly less than one reaction time. Because feedback from the periphery is too slow to cause these corrections, it was reasoned that subjects detected their own errors before the movement was begun through monitoring of efference copy.

Efference Copy in the Schema Theory

The present theory sees efference copy as necessary to provide two kinds of information. First, the subject must know that he has executed a motor command so that the information he receives from the periphery can be interpreted as resulting from active rather than passive movement. Without being at all specific about the details, we assume the existence of some central mechanism that records that a motor program has been run off; no literal copy of the muscle commands is assumed to be stored. Second, the notion of efference as a feed-forward process is assumed in the present theory. Before the movement, the expected sensory consequences are generated from the recognition schema, and this information is fed forward so that the incoming response-produced exteroceptive feedback and proprioceptive feedback can be evaluated properly, resulting in the detection of error. In addition, the motor program output is thought to contain gamma efferent signals that are fed forward to the muscle spindles, preparing the local musculature for reflex-based corrections that ensure faithful execution of the program. Further involvements of efference copy are not proposed in the present theory. The evidence cited above apparently supporting the outflow model (i.e., learning with deafferentation and rapid error correction) can be handled with the present theoretical structure.

In the deafferentation studies, the monkeys had all that was necessary to learn the avoidance response without efference. When the shock came on, the animals began a series of random movements in an attempt to escape. Eventually, quite by accident, the animal squeezed the bulb and the shock ceased. According to the present theory, at this time the animal had stored both the response specifications (for the random movements ordered) and the actual outcome (the cessation of the shock). Which specification was associated with shock cessation was not clear, however, since many responses were being ordered at nearly the same time, but the schema rule notes the relationship between the various response specifications and

the cessation of the shock. On the next shock trial, the many response specifications related to the desired outcome were executed, and eventually the hand squeeze occurred again, strengthening the relationship between the response specifications and the outcome. Hence, the animal could learn without direct feedback from the limb, because he had the essential ingredients on each trial to form the schema rule: the actual outcome and the response specifications. The logic presented here is similar to Guthrie's (1952) notion of stimulus-response contiguity. The notion of efference copy as a central feedback loop indicating the limb movements is not necessary to explain these findings.

The error correction studies can also be explained by the present theory without involving efference. In these studies, the reaction times for correct responses (the interval from stimulus until first movement) was greater than for incorrect responses, suggesting that subjects were anticipating on error trials. Theoretically, on an error trial the subject incorrectly guesses the direction of the upcoming movement and initiates response specifications in advance, with actual movement not beginning until after the stimulus. Before the program is initiated, the subject generates the expected sensory consequences, which include the expected (if the guess is correct) vision from the stimulus array. When the (unexpected) stimulus comes on, the incorrect program is already being run off when the subject receives a mismatch between the expected and actual sensory consequences. When the mismatch occurs, a visually based correction begins and the program for the correct movement is initiated. However, because the first (incorrect) program must run its course for one reaction time or longer, the beginnings of the incorrect movement occur before the correction. This explanation predicts that the time from the stimulus until the first movement in the correct direction should approximate one reaction time. These values were 367 msec (Angel & Higgins, 1969), 292 msec (Megaw, 1972), and 285 msec in a recent study by Gordon (1975).

all of which are well within the normal range of reaction time.

Thus, the present theory uses efference copy as a mechanism that informs the subject that a motor program has been initiated, and represents various feed-forward processes. Feeding forward the expected proprioceptive and exteroceptive feedback is essential for the accurate detection of errors, and the generation of gamma efferent signals readies the musculature for subsequent action. There is no place in the theory, as presently stated, for efference copy as a mechanism that can allow the subject to determine errors in terms of reaching the desired goal, because the current evidence can be explained without resorting to such a notion.

POSSIBLE FUTURE DIRECTIONS

While the present theory can solve a number of problems with existing theoretical points of view and has strong empirical support in some areas, the evidence for it is weak or generally lacking in other areas. There are numerous directions in which one could move, including asking about individual differences in schema formation, how developmental influences are manifested, and how large a given schema might be in terms of the range of movements encompassed by it. But the answers to some essential questions should be sought first to determine whether the theoretical ideas presented here are worthy of further work. This section deals with some of these questions.

First, we need stronger evidence for motor recall and motor recognition schemata. We might take comfort that the schema notion seems acceptable enough for the visual pattern-recognition literature, but motor skills researchers have been embarrassed before by taking the work in verbal learning and applying it to motor behavior without a serious research effort to determine whether the application is reasonable. A number of methods of testing the schema notion for motor behavior have been mentioned, including evidence that transfer from a variety of experiences to a new instance of the same class is high, that increased variation in practice should increase transfer, and that responses

presumed to be governed by the same motor program share in temporal and spatial organization (e.g., Glencross, 1973). The notion of the schema is clearly essential to the theory, and it seems logical that the credibility of this idea should be strongly tested initially.

It is possible that it will be difficult to demonstrate schema learning in adult subjects, and perhaps a more profitable population would be children. Many researchers adhere to the view that most of motor learning occurs in children and that adult learning is merely a recombination of the old habits; it would therefore make sense to study the acquisition of these new skills at the age when they develop most rapidly. Failing to demonstrate schema learning in adults does not necessarily mean that schemata do not exist, since the schemata for certain classes of highly practical activity (such as throwing) could have been developed years earlier.

Another important prediction made by the theory is the possibility of learning rapid tasks without KR after some initial KR practice. This has not been shown clearly so far, and additional work should be conducted with longer KR practice periods, with tasks in which complexity is higher than in the simple slide-movement tasks, and with more extended no-KR practice periods. In addition, the predictions about tasks in which accuracy is not a goal, as well as the predictions referring to open and closed tasks, need to be tested.

A final word concerns potential applicability. While the theory was certainly not developed for application to the teaching of motor skills, the predictions seem to have a great deal in common with the notion of movement education. While there are many branches of this idea, one of the important thrusts is that children should engage in activities stressing variety in movement patterns (e.g., jump over an object in as many ways as possible). Movement education people do not talk of schemata, of course, but one interpretation of their method is that it develops schemata in children. These schemata will then be useful in an open-skills situation where a jumping response is re-

quired. This area has been devoid of theory for a long time, and perhaps the present views can be of assistance.

REFERENCE NOTES

1. Schmidt, R. A. *Open- and closed-loop processes in discrete motor tasks: A review and theory*. Paper presented at the national convention of the North American Society for the Psychology of Sport and Physical Activity, Anaheim, Calif., 1974.
2. Schmidt, R. A., & McCabe, J. *Changes in motor program utilization over extended practice*. Unpublished manuscript, 1973. (Available from R. A. Schmidt, Department of Physical Education, University of Southern California, Los Angeles, California 90007.)
3. Armstrong, T. R. *Training for the production of memorized movement patterns* (Tech. Rep. 26). Ann Arbor, Mich.: University of Michigan, Human Performance Center, August, 1970.

REFERENCES

- Adams, J. A. Motor skills. *Annual Review of Psychology*, 1964, 15, 181-202.
- Adams, J. A. A closed-loop theory of motor learning. *Journal of Motor Behavior*, 1971, 3, 111-150.
- Adams, J. A., & Dijkstra, S. Short-term memory for motor responses. *Journal of Experimental Psychology*, 1966, 71, 314-318.
- Adams, J. A., & Goetz, E. T. Feedback and practice as variables in error detection and correction. *Journal of Motor Behavior*, 1973, 5, 217-224.
- Adams, J. A., Goetz, E. T., & Marshall, P. H. Response feedback and motor learning. *Journal of Experimental Psychology*, 1972, 92, 291-397.
- Angel, R. W., Garland, H., & Fischler, M. Tracking errors amended without visual feedback. *Journal of Experimental Psychology*, 1971, 89, 422-424.
- Angel, R. W., & Higgins, J. R. Correction of false moves in pursuit tracking. *Journal of Experimental Psychology*, 1969, 82, 185-187.
- Anokhin, P. K. Cybernetics and the integrative activity of the brain. In M. Cole & I. Maltzman (Eds.), *A handbook of contemporary Soviet psychology*. New York: Basic Books, 1969.
- Bahrack, H. P. An analysis of stimulus variables influencing the proprioceptive control of movements. *Psychological Review*, 1957, 64, 324-328.
- Bartlett, F. C. *Remembering*. Cambridge, England: Cambridge University Press, 1932.
- Bereiter, C. Some persisting dilemmas in the measurement of change. In C. W. Harris (Ed.), *Problems in measuring change*. Madison: University of Wisconsin Press, 1963.
- Bernstein, N. *The co-ordination and regulation of movements*. London: Pergamon Press, 1967.
- Bilodeau, E. A., & Bilodeau, I. M. Variable frequency of knowledge of results and the learning of a simple skill. *Journal of Experimental Psychology*, 1958, 55, 379-383.
- Bilodeau, E. A., Bilodeau, I. M., & Schumsky, D. A. Some effects of introducing and withdrawing knowledge of results early and late in practice. *Journal of Experimental Psychology*, 1959, 58, 142-144.
- Bilodeau, I. M. Information feedback. In E. A. Bilodeau (Ed.), *Acquisition of skill*. New York: Academic Press, 1966.
- Broer, M. *Efficiency of human movement*. Philadelphia, Pa.: Saunders, 1973.
- Brooks, V. B. Some examples of programmed limb movements. *Brain Research*, 1974, 71, 299-308.
- Chang, H. T. Activation of internuncial neurons through collaterals of pyramidal fibers at cortical level. *Journal of Neurophysiology*, 1955, 18, 452-471.
- Chernikoff, R., & Taylor, F. V. Reaction time to kinesthetic stimulation resulting from sudden arm displacement. *Journal of Experimental Psychology*, 1952, 43, 1-8.
- Dobbins, D. A., & Rarick, G. L. Intra-individual variability of normal and educable mentally retarded boys on tests of throwing accuracy. *Research Quarterly*, in press.
- Duncan, D. P. Transfer after training with single versus multiple tasks. *Journal of Experimental Psychology*, 1958, 55, 63-72.
- Easton, T. A. On the normal use of reflexes. *American Scientist*, 1972, 60, 591-599.
- Edmonds, E. M., & Evans, S. H. Schema learning without a prototype. *Psychonomic Science*, 1966, 5, 247-248.
- Edmonds, E. M., Evans, S. H., & Mueller, M. R. Learning how to learn schemata. *Psychonomic Science*, 1966, 6, 177-178.
- Edmonds, E. M., & Mueller, M. R. The role of schemata in perceptual learning. *Psychonomic Science*, 1966, 8, 230.
- Edmonds, E. M., Mueller, M. R., & Evans, S. H. Effects of knowledge of results on mixed schema discrimination. *Psychonomic Science*, 1966, 6, 377-378.
- Ellis, H. C. *The transfer of learning*. New York: Macmillan, 1965.
- Ells, J. G. *Attentional requirements of movement control*. Unpublished doctoral dissertation, University of Oregon, 1969.
- Evans, S. H. Redundancy as a variable in pattern recognition. *Psychological Bulletin*, 1967, 67, 104-113. (a)
- Evans, S. H. A brief statement of schema theory. *Psychonomic Science*, 1967, 8, 87-88. (b)
- Evans, S. H., & Edmonds, E. M. Schema discrimination as a function of training. *Psychonomic Science*, 1966, 5, 247-248.
- Festinger, I., & Canon, L. K. Information about spatial location based on knowledge about efferecence. *Psychological Review*, 1965, 72, 378-384.
- Fitts, P. M. The information capacity of the human motor system in controlling the amplitude

- of movement. *Journal of Experimental Psychology*, 1954, 47, 381-391.
- Gentile, A. M. A working model of skill acquisition with application to teaching. *Quest*, 1972, 17, 3-23.
- Glencross, D. J. The effects of changes in direction, load, and amplitude of movement on gradation of effort. *Journal of Motor Behavior*, 1973, 5, 207-216.
- Gordon, G. B. *Correction of errors in discrete step tracking*. Unpublished masters thesis, University of Michigan, 1975.
- Granit, R. *The basis of motor control*. New York: Academic Press, 1970.
- Guthrie, E. R. *The psychology of learning*. New York: Harper, 1952.
- Gyr, J. W. Is a theory of direct visual perception adequate? *Psychological Bulletin*, 1972, 77, 246-261.
- Head, H. *Aphasia and kindred disorders of speech*. Cambridge, England: Cambridge University Press, 1926.
- Helmholtz, H. v. *Treatise on physiological optics* (Vol. 3, 3rd ed., J. P. C. Southall, Ed. and Trans.). Menasha, Wis.: Optical Society of America, 1925.
- Henderson, S. *The role of feedback in the development and maintenance of a complex motor skill*. Unpublished doctoral thesis, University of Waterloo, 1974.
- Henry, F. M. Increased response latency for complicated movements and a "memory-drum" theory of neuromotor reaction. *Research Quarterly*, 1960, 31, 448-458.
- Henry, F. M. Constant and variable performance errors within a group of individuals. *Journal of Motor Behavior*, 1974, 6, 149-154.
- Higgins, J. R., & Angel, R. W. Correction of tracking errors without sensory feedback. *Journal of Experimental Psychology*, 1970, 84, 412-416.
- Higgins, J. R., & Spaeth, R. K. Relationship between consistency of movement and environmental condition. *Quest*, 1972, 17, 61-69.
- Holst, E. v. Relations between the central nervous system and the peripheral organs. *British Journal of Animal Behavior*, 1954, 2, 89-94.
- Hull, C. L. *Principles of behavior*. New York: Appleton Century, 1943.
- Jones, B. Is proprioception important for skilled performance? *Journal of Motor Behavior*, 1974, 6, 33-45.
- Keele, S. W. Movement control in skilled motor performance. *Psychological Bulletin*, 1968, 70, 387-403.
- Keele, S. W., & Posner, M. I. Processing of feedback in rapid movements. *Journal of Experimental Psychology*, 1968, 77, 353-363.
- Konorski, J. *Integrative activity of the brain*. Chicago: University of Chicago Press, 1967.
- Lachman, R. The model in theory construction. *Psychological Review*, 1960, 67, 113-129.
- Lashley, K. S. The accuracy of movement in the absence of excitation from the moving organ. *American Journal of Physiology*, 1917, 43, 169-194.
- Laszlo, J. I. Training of fast tapping with reduction of kinesthetic, tactile, visual, and auditory sensations. *Quarterly Journal of Experimental Psychology*, 1967, 19, 344-349.
- Laszlo, J. I., & Bairstow, P. J. The compression block technique: A note on procedure. *Journal of Motor Behavior*, 1971, 3, 313-317.
- Li, C. L. Activity of interneurons in the motor cortex. In H. H. Jasper, L. D. Procter, R. S. Knighton, W. C. Norsbay, & R. T. Costello (Eds.), *Reticular formation of the brain*. Boston: Little, Brown, 1958.
- MacNeilage, P. F. Motor control of serial ordering of speech. *Psychological Review*, 1970, 77, 182-196.
- MacNeilage, P. F., & MacNeilage, L. A. Central processes controlling speech production during sleep and waking. In F. J. McGuigan (Ed.), *The psychophysiology of thinking*. New York: Academic Press, 1973.
- Marshall, P. H. Recognition and recall in short-term motor memory. *Journal of Experimental Psychology*, 1972, 95, 147-153.
- Megaw, E. D. Directional errors and their correction in a discrete tracking task. *Ergonomics*, 1972, 15, 633-643.
- Merton, P. A. How we control the contraction of our muscles. *Scientific American*, 1972, 225, 3-10.
- Newell, K. M. Knowledge of results and motor learning. *Journal of Motor Behavior*, 1974, 6, 235-244.
- Newell, K. M., & Boucher, J. P. Motor response recognition: Two processes. *Journal of Motor Behavior*, 1974, 6, 81-86.
- Newell, K. M., & Chew, R. A. Recall and recognition in motor learning. *Journal of Motor Behavior*, 1974, 6, 245-253.
- Nottebohm, F. The ontogeny of bird song. *Science*, 1970, 167, 950-956.
- Pew, R. W. Acquisition of hierarchical control over the temporal organization of a skill. *Journal of Experimental Psychology*, 1966, 71, 764-771.
- Pew, R. W. Human perceptual-motor performance. In B. H. Kantowitz (Ed.), *Human information processing: Tutorials in performance and cognition*. New York: Erlbaum, 1974.
- Posner, M. I., & Keele, S. W. On the genesis of abstract ideas. *Journal of Experimental Psychology*, 1968, 77, 353-363.
- Posner, M. I., & Keele, S. W. Retention of abstract ideas. *Journal of Experimental Psychology*, 1970, 83, 304-308.
- Poulton, E. C. On prediction in skilled movements. *Psychological Bulletin*, 1957, 54, 467-473.
- Schmidt, R. A. The case against learning and forgetting scores. *Journal of Motor Behavior*, 1972, 4, 79-88. (a)
- Schmidt, R. A. The index of preprogramming (IP): A statistical method for evaluating the

- role of feedback in simple movements. *Psychonomic Science*, 1972, 27, 83-85. (b)
- Schmidt, R. A. A reaction to Ron Marteniuk's paper. In M. G. Wade & R. Martens (Eds.), *Psychology of motor behavior and sport*. Urbana, Ill.: Human Kinetics Publishers, 1974.
- Schmidt, R. A. *Motor skills*. New York: Harper & Row, 1975.
- Schmidt, R. A. The schema as a solution to some persistent problems in motor-learning theory. In G. E. Stelmach (Ed.), *Motor control: Issues and trends*. New York: Academic Press, in press.
- Schmidt, R. A., Christenson, R., & Rogers, P. Some evidence for the independence of recall and recognition in motor behavior. *Journal of Motor Behavior*, in press.
- Schmidt, R. A., & Russell, D. G. Movement velocity and movement time as determiners of the degree of preprogramming in simple movements. *Journal of Experimental Psychology*, 1972, 96, 315-320.
- Schmidt, R. A., & White, J. L. Evidence for an error detection mechanism in motor skills: A test of Adams' closed-loop theory. *Journal of Motor Behavior*, 1972, 4, 143-153.
- Schmidt, R. A., & Wrisberg, C. A. Further tests of the Adams' closed-loop theory: Response-produced feedback and the error-detection mechanism. *Journal of Motor Behavior*, 1973, 3, 155-164.
- Schutz, R. W., & Roy, E. A. Absolute error: The devil in disguise. *Journal of Motor Behavior*, 1973, 5, 141-154.
- Sokolov, E. N. The modeling properties of the nervous system. In M. Cole & I. Maltzman (Eds.), *A handbook of contemporary Soviet psychology*. New York: Basic Books, 1969.
- Solley, C. M. Reduction of error with practice in perception of the postural vertical. *Journal of Experimental Psychology*, 1956, 52, 329-333.
- Taub, E., & Berman, A. J. Movement and learning in the absence of sensory feedback. In S. J. Freedman (Ed.), *The neuropsychology of spatially oriented behavior*. Homewood, Ill.: Dorsey Press, 1968.
- Underwood, B. J. Are we overloading memory? In A. W. Melton & E. Martin (Eds.), *Coding processes in human memory*. Washington, D.C.: V. H. Winston, 1972.
- Williams, I. D., & Rodney, M. Intrinsic feedback, interpolation, and the closed-loop theory. *Journal of Motor Behavior*, in press.
- Wilson, D. M. The central nervous control of flight in a locust. *Journal of Experimental Biology*, 1961, 38, 471-490.
- Woodworth, R. S. *Experimental psychology*. New York: Holt, 1938.
- Wrisberg, C. A., & Schmidt, R. A. A note on improved motor performance without knowledge of results. *Journal of Motor Behavior*, in press.

(Received October 14, 1974)

Structure of Memory Traces

Endel Tulving

University of Toronto, Toronto, Canada,
and Yale University

Michael J. Watkins
Yale University

A new descriptive theory of memory traces is presented. The memory trace is (a) conceptualized as a collection of trace elements and (b) defined in terms of the relation between the conditions and the product of retrieval. The properties of a trace thus defined are quantitatively described by measuring the gross, common, and reduced valences of two (or more) retrieval cues. These valences are determined by successively probing the target event with each of the cues. The data yielded by the successive probes are then used to construct the matrix or structure of the trace by means of the *reduction method*. The logic of this method, and hence the general theory, is applicable to a large variety of to-be-remembered material. A demonstration experiment showed that the structure of the traces of to-be-remembered word-events is sensitive to the conditions of initial encoding, and that forgetting of these events, under conditions of output interference, consists in a distinctive change in the pattern of trace elements. Some potential criticisms of the theory are considered.

Memory trace refers to a change in the memory system that results from the perception and encoding of an event. It constitutes a necessary condition for the subsequent retrieval of information about the event. A long-standing problem in the study of memory concerns the nature and properties of memory traces.

In a typical laboratory situation a person sees or hears a familiar word, which he is asked to store in memory. The perception of the word-event produces a change in the system, that is, it creates a memory trace. How do we conceptualize the trace of the word-event? How do we define it? How do we describe it quantitatively? And how do we relate our conceptualization, definition, and description of the trace to other ideas about memory?

This article suggests some possible answers to these questions by proposing a simple descriptive theory of the memory trace. Traces are (a) conceptualized as structured aggregates of trace elements, (b)

defined as relations between the conditions and the products of retrieval, and (c) quantitatively described in terms of valences (effectiveness) of retrieval cues and of the interrelations of cue valences. The theory is embodied in a set of experimental and analytic procedures for specifying the structure of a memory trace.

Our conceptualization of memory traces as aggregates of elements is in keeping with currently popular notions of concepts and words being represented in memory as bundles of attributes or features (e.g., Anisfeld & Knapp, 1968; Bower, 1967; Morton, 1969; Underwood, 1969). We discuss trace elements, rather than features and attributes, solely for convenience in terminology.

We will not attempt to relate our trace-element approach to the other major theoretical tradition in characterizing memory representation, namely the associative network approach (e.g., Anderson & Bower, 1973; Kintsch, 1974; Rumelhart, Lindsay, & Norman, 1972). Our objective here is not to present a comprehensive review and analysis of the accomplishments of the various theories of memory representation; rather, it is to describe and discuss one possible approach to the problem of specifying properties of memory traces.

This research was supported by National Science Foundation Grant GB-40208X, and National Research Council of Canada Grant A8632.

Michael J. Watkins is now at the University of Toronto.

Requests for reprints should be sent to Endel Tulving, Department of Psychology, University of Toronto, Toronto M5S 1A1, Ontario, Canada.

TRACES AS RELATIONS BETWEEN OBSERVABLES

Since memory traces cannot be observed directly, they must be described indirectly, that is, in terms of things that *can* be observed. Three classes of observable entities are always involved in remembering: (a) the input into the memory system, (b) the retrieval query or cue, and (c) the output from the system. The relation between the input and the output was for a long time considered the only source of information about the memory trace. For instance, the correlation between some variable *X* (e.g., frequency, recency, etc.) and the probability of recall or recognition has been regarded as evidence that trace strength is a function of variable *X*. Thus, a property of the trace, its strength, has been defined in terms of the input-output relation.

There are at least three difficulties with the idea of traces as input-output relations. First, it encourages the notion of the trace as a copy of what has been presented to the memory system. The copy theory, even if rendered more plausible by assuming that the trace is but an incomplete or impoverished counterpart of the original object or event, seems intuitively unreasonable.

Second, a definition of the trace as an input-output relation cannot cope with the fact that the output from the system varies not only with the input but also with the conditions under which retrieval is attempted. This difficulty is usually avoided by observing the consequences of input only under invariant retrieval conditions. Unfortunately, this strategy precludes the study of some of the most interesting aspects of memory.

The third weakness of the input-output relation as a basis for defining traces is perhaps the most important one: It precludes making psychologically meaningful and empirically verifiable statements about the relations between input conditions and the trace of an item. If the trace is defined as a relation between input and output, its characterization must include a description of input conditions. Therefore, the relation

between input conditions and the trace is a matter of logic, rather than experiment.

What we would like to have is some way of defining traces that would (a) permit us to arrive at a picture of the memory of an input event that is not a copy of the event, (b) reconcile the idea of a fixed trace with the fact that its activation will depend on the retrieval environment, and (c) foster questions about the relation between input conditions and the trace of an item whose nominal identity is held constant.

These objectives can be achieved if memory traces are defined in terms of the relation between conditions and products of retrieval, or as "a set of question-output relations" (Tulving & Bower, 1974, p. 294). The basic ideas are simple. Rather than thinking of the trace only as an aftereffect of the initial registration of an event, we think of it as a "cognitive blueprint" that specifies the conditions under which the recollection of the event will occur. Then, by assuming the encoding specificity principle (Tulving & Thomson, 1973)—the principle that any retrieval cue is effective only to the extent that its informational contents match the information contained in the trace of the event—we can interpret the observed effectiveness of different retrieval cues as evidence about the properties of the trace. Thus the definition of the trace in terms of the relation between the queries directed at the system and the output from the system leads to the description of the trace properties in terms of the effectiveness of various retrieval cues in producing recall.

Such an approach abates some of the difficulties inherent in the definition of memory traces as input-output relations. The description of the trace can emerge as something other than a copy of the input, since effective retrieval cues may be qualitatively quite different from the original event. Also, there would be no logical paradox about the fixed traces having variable manifestations, since the logic of the approach includes the assumption that the output from the system depends on the retrieval information (the cue) as well as on the stored information (the trace). Finally, this approach encourages questions about relations between input

conditions and the resulting traces of one and the same item.

The major part of this article consists of the elaboration and illustration of this approach. But we will first briefly consider two currently popular approaches to the problem of describing memory traces.

TRACES AS FEATURE BUNDLES: TWO EXTANT METHODS

Many different methods have been used by various investigators in attempts to get at the properties of memory traces conceptualized as feature bundles. In a recent review, Tulving and Bower (1974) discussed 11 methods, and interested readers may wish to look at that review for a fuller background for the present discussion. We will mention here only two of the methods—those with which ours might be most directly contrasted.

One is Wickens's (1970, 1972) method of attribute shifts, which involves giving three or four successive short-term recall tests with the Brown-Peterson technique, with the study items sharing some particular perceptual or conceptual feature. Recall is typically found to decrease rapidly across the successive tests. A "critical" test is then given, in which the study items do not share one or more attributes that the items of previous tests had in common. The critical test often yields a high level of recall, compared with that for a control test in which no attribute change takes place. Wickens has argued that such a shift effect indicates that the systematically varied features were encoded into memory.

Although Wickens's attribute-shift method has produced a lot of interesting data, it does have a number of problems. One is that it appears to have quite limited applicability. So far, it has been used only within the confines of the conventional Brown-Peterson paradigm and under rather specific conditions of retention and intertest intervals. It is not known whether the method could be extended to measure the composition of traces of other sorts of input or with intertraces of other sorts of input or with intertraces measured in minutes, hours, or years. Another difficulty has to do with the fact that it is sometimes necessary to provide

specific information to the subject before the attribute-shift effect is obtained—the change in the attribute is not always sufficient (Gardiner, Craik, & Birtwistle, 1972). Thus we are faced with the difficulty of deciding exactly what kind of retrieval information should be provided to identify trace features. Still another shortcoming of the method lies in the fact that the investigation of each potential encoding feature involves the use of a special set of items, a set which, as Underwood (1972) has pointed out, could serve to "prime" the feature under investigation. It follows that we cannot be certain that the dimensions of encoding suggested in the attribute-shift data have any generality beyond the precise conditions of the paradigm. Finally, the attribute-shift method provides no easy solution to the problem of determining the extent to which any two types of attribute shifts reflect the operation of a single underlying process. This is an aspect of the problem of establishing the relative dependence of trace features, a problem that occurs in many extant methods of describing traces (see Tulving & Bower, 1974).

Wickens's attribute-shift method is based on the implicit definition of the trace as an input-output relation. Although the input conditions in the method are specified in the form of differences between sets of to-be-remembered materials, and the output is indexed by a derived measure involving a comparison between two conditions, the effects of an attribute shift, and more particularly the ranking of the relative effectiveness of several attribute shifts (e.g., Wickens, 1970, Fig. 7; 1972, Fig. 3), do provide a picture of the trace as a relation between input and output. And, as we have seen, such a formulation has serious limitations.

The other method with which the theory and technique to be developed in this article can be contrasted is the method of false positive recognition errors. It was first described by Underwood (1965) and has been used as a tool for identifying component features of memory traces by Anisfeld and Knapp (1968) as well as many others (e.g., Bach & Underwood, 1970; Elias & Perfetti, 1973; Kosslyn & Bower, 1974). A person studies

a list of items (words, sentences, or the like), and later takes a recognition test. Beside copies of study items and unrelated distractors, the test includes items that share particular features with study items. Thus, if one of the study items was the word CHAIR, the test might include the semantically related word *table* and the acoustically related word *hare*. To the extent that the person checks these distractor items as "old" more frequently than unrelated distractors, the features common to the study and the distractor items are implicated as a part of the trace of the study item. The features are usually designated by such terms as "associated words" or "rhyming words."

The method of false positive recognition errors is based on an implicit definition of the trace as a query-output relation. It is therefore free of the special difficulties inherent in input-output definitions of the trace. But it does have certain other difficulties. For instance, the experimenter gains information about trace features only to the extent that subjects make false positive "errors." Hence, the evidence for trace features depends, among other things, on the factors determining the frequency of such errors; yet some of these factors, such as the ratio of target to distractor items, probably have little to do with the nature of the trace. An attendant practical problem is that under typical conditions, where the subjects strive for accuracy, the density of the data that can be used to make inferences about trace properties is quite low, necessitating the testing of many subjects for stable results.

A second problem with the method is that the experimenter does not always know why the subject checks a particular distractor item as "old." If the study list contains the word CHAIR and the subject checks the distractor word *hare*, he may do so because of acoustic confusion with CHAIR, but he may also do it for some other reason, perhaps because of an idiosyncratically meaningful semantic confusion with some other study word. The very fact that control words in the test, selected to be unrelated to the study items, are checked as "old" suggests that the experimenter cannot always identify the basis for a false positive response.

A third problem has been discussed by Tulving and Bower (1974, p. 280) and will be only briefly mentioned here. Inferences about trace features require assumptions about features lost from the trace rather than features present in the test item. If the subject checks the distractor word *table* after having seen CHAIR in the list, he presumably does so because of an overlap of semantic features between *table* and CHAIR. But this overlap should only produce an "old" response to the extent that other features that normally distinguish *table* and CHAIR have been lost from the trace of CHAIR or are not extracted from the test stimulus *table*. It is not clear what kind of theory is needed here to relate the empirical data to the hypothetical memory trace.

Finally, there is again the problem concerning the relations among features, their independence and separability. Assume that the word CHAIR has been presented as a to-be-remembered item and that in the recognition test two distractors related to CHAIR, *table* and *sit*, are both checked as false positives. Does this mean that the trace of CHAIR includes one feature (e.g., "associative") or two features (e.g., "furniture" and "activity-instrument")? To rely on the experimenter's intuitions about the relation between two types of distractor items is not very satisfactory. Even where two items may appear to be very different (as with a semantically related word such as *table* and an acoustically related word such as *hare*), there is little to justify the conclusion that they have *nothing* in common. Evidence about the relations between different types of distractor items should be derived with an objective procedure, and so far no such procedure has been proposed.

RETRIEVAL CUING AND THE REDUCTION METHOD

Retrieval cuing as a method for describing memory traces has been briefly discussed by Tulving and Bower (1974). It has been used as a tool for gaining information about the composition of memory traces in a number of experiments (e.g., Barclay, Bransford, Franks, McCarrell, & Nitsch, 1974; Ghatala & Hurlbut, 1973; Light, 1972; Nel-

son & Brooks, 1974), primarily by providing a list of possible trace features or attributes. The present article, in essence, describes an extension of the method of retrieval cuing. The extension consists in the *successive* probing of a given trace with two or more different types of retrieval cues. By subjecting the results of such a procedure to a form of analysis that we call the *reduction method*, the relations among the retrieval cues with respect to the trace may be observed. And in the context of our theory, the observed cue relationships describe the structure of the referent trace.

Central to our approach is the encoding specificity principle mentioned earlier: A retrieval cue is effective to the extent that its informational contents match the informational contents of the trace. Thus we can infer informational properties of a particular trace and quantitatively specify its constituent trace elements by measuring the effectiveness of different retrieval cues.

At this point it is useful to introduce the idea of the valence of a cue, or more precisely the valence of a cue with respect to some particular trace. In fact we will define three types of valence: *gross valence* (often abbreviated to *valence*), *reduced valence*, and *common valence*. The gross valence of a cue with respect to the trace of a certain event refers simply to the probability with which that event can be recalled in the presence of the cue. In typical experimental situations where the subject's task is to reproduce the name of a word-event, gross valence refers to the probability of recall of the word in the presence of the cue. The valence will therefore assume a value between zero and unity. The valence of a cue must always be specified, at least implicitly, with reference to some defined trace, for the valence of a cue will vary widely for traces of different types of event—even among events involving the presentation of the same item under different conditions. Reduced valence refers to a cue valence that has been reduced by some other cue. The valence of cue *X* reduced by cue *Y* is given by the probability that the target can be recalled to cue *X* and cannot be recalled to cue *Y*. The common valence of two cues, such as cue *X*

and cue *Y*, refers to the probability that the trace can be retrieved by cue *X* as well as by cue *Y*. It follows that the gross valence of cue *X* equals the sum of the reduced (by cue *Y*) valence of cue *X* and the common valence of cues *X* and *Y*.

Successive Probes

Consider now a simple experimental situation in which a subject is shown a familiar word among others in a to-be-remembered list. After the presentation of the list the subject is given two temporally separated cues for the target word in question, and asked to use the cues in recalling the target word. For instance, if the target word is CHAIR, the subject could be given the extralist cue *table*, perhaps with a statement of the conceptual relation of the cue to the target word. Thus, the complete cue might be "*table* (associated with) _____." Some time later in the test the subject is given another cue, such as "*hare* (rhymes with) _____." Of course, additional probes for the same target word could also be given, but we will consider only the situation in which two probes are used.

If these simple operations of presenting a word and probing it with two successive cues are replicated many times, frequencies of various possible outcomes can be used to estimate the valences of the cues. To do so, the data are summarized in the form of a 2×2 contingency table. One such data matrix from an imaginary experiment is shown in Table 1. Here one cue, *X*, was always presented as the first probe, and another cue, *Y*, as the second. The data in Table 1 tell us that cue *X* effected recall of the target item 45% of the time, and cue *Y* 60% of the time. We can also see that 40% of the time the target word was recalled to both cues, 5% of the time only to the first cue, and 20% of the time only to the second cue. Using our terminology, we can say that the gross valence of cue *X* was .45 and of cue *Y* was .60, that the common valence of cue *X* and cue *Y* was .40, and that the valence of cue *X* reduced by cue *Y* was .05, while the reduced valence of cue *Y* was .20.

In keeping with the encoding specificity principle, we regard the pattern of cue va-

TABLE 1

GROSS, COMMON, AND REDUCED VALENCES OF TWO CUES: HYPOTHETICAL DATA

First cue X	Second cue Y		Total
	+	-	
+	.40	.05	.45
-	.20	.35	.55
Total	.60	.40	

lences shown in Table 1 as describing the trace of the target item. In other words, we have described the trace here in terms of the gross, common, and reduced valences of the two cues. Since the common and reduced valences are not predictable from the gross valences, the whole pattern of data tells us more about the memory trace than do the gross valences alone. Indeed, we might say that the pattern of the two valences tells us something about the *structure* of the memory trace.

To formalize the distinction between valences of cues and the properties of the trace, we will refer to components of the trace as elements and define them in terms of retrieval cues. Thus, Table 1 can be said to depict a trace containing 40% XY elements, 5% of $X\bar{Y}$ elements, 20% of $\bar{X}Y$ elements, and 35% $\bar{X}\bar{Y}$ elements. It is easy to imagine a situation in which the gross valences of cue X and cue Y are the same as those shown in Table 1, but the cell entries in the matrix are quite different. We would then say that even though the two traces had certain properties in common, their informational structures were different.

Reduction Method

Although the basic idea underlying the method of retrieval cuing for quantitatively describing memory traces is simple, its implementation is a bit more complicated. Complications arise because the valences of the two cues, X and Y , may depend on the order in which they are presented. This would be expected if the first cue changed the trace (retrieval-induced recoding of the trace), since the two cues would then be applied to different traces. If it were not for such potential confounding, we could let the data pattern in Table 1 represent the

structure of the trace. But given that it is indeed possible, if not probable, that prompting recall of an event with one cue will change the valence of a second cue, another method for describing the trace is required. Our response to this requirement has been the development of the *reduction method*.

Consider again the basic experimental situation that we have described: A trace of a certain list item exists in memory and is probed successively with two cues, X and Y . But now the two cues are applied to the trace in two different cue-order conditions. In one condition the target is first probed with cue X and then, after other interpolated tests, with cue Y ; in the other condition the order of the two cues is reversed. Data from the two cue-order conditions are tabulated separately in the form of two 2×2 tables. The data from these two tables are then combined into a third table, one that directly describes the trace structure.

Table 2 shows a set of data from an imaginary experiment in which gross, common, and reduced valences of two cues, X and Y , are obtained in two cue-order conditions as just described. The 2×2 table on the left (identical to Table 1) shows the results for the condition in which cue X was presented first and cue Y second, and the middle 2×2 table depicts the results with cue Y presented first and cue X second. We will refer to these tables as *data matrices*. The right-hand 2×2 table, which we will call the *trace matrix*, is composed of selected data from the two data matrices. As Table 2 shows, only two entries in each data matrix are used in constructing the trace matrix: the gross valence of the first cue and the reduced valence of the second. The remaining entries in the trace matrix are arrived at by subtraction, in keeping with the properties of 2×2 contingency tables.

The trace matrix is constructed in a way that minimizes the effect of cuing order and possible retrieval-induced recoding of the trace. Our basic assumption is that presenting a cue might cause a recoding of the target trace only if the cue is successful in effecting recall; a cue that does not elicit recall of the target item is assumed to leave the trace intact. Thus, in describing the

TABLE 2
TWO DATA MATRICES AND RESULTANT TRACE MATRIX: HYPOTHETICAL DATA

Data matrices							Trace matrix				
First cue X	Second cue Y			First cue Y	Second cue X			Cue X	Cue Y		
	+	-	Total		+	-	Total		+	-	Total
+	.40	.05	.45	+	.27	.25	.52	+	.32	.13	.45
-	.20	.35	.55	-	.13	.35	.48	-	.20	.35	.55
Total	.60	.40		Total	.40	.60		Total	.52	.48	

Note. Boldface indicates two entries from each data matrix (gross valence of first cue; reduced valence of second cue) used in constructing trace matrix.

trace we must take care not to use data that may be contaminated by the effects of a successful first cue. As Table 2 shows, no such contaminated data from the data matrices are used in constructing the trace matrix. The gross valence of the first cue is, of course, free of any contamination by the act of retrieval, and so is the reduced valence of the second cue, since this is based on only those cases in which the first cue did not effect recall.¹

Two things should be noted about this method of estimating the gross, common, and reduced valences of the two cues with respect to a given trace. First, since we wish to avoid the potential problem of a change in the structure of a trace as a consequence of the act of retrieval, and since we therefore must exclude the data based on the retrieval of items previously retrieved, it is necessary to derive a trace matrix from two separate data matrices, each corresponding to one of the two possible test sequences. Provided our assumption is correct, this procedure gets around the difficulty created by the possible effects of retrieval on the trace. The effects of retrieving a trace on one occasion on retrieving it on a subsequent

occasion represent an interesting research problem, but whatever these effects are they are not relevant to the reduction method.

The second point has to do with the problem of combining the data from the two data matrices into a single trace matrix. In the illustration shown in Table 2, the critical values for the two data matrices were entered directly into the trace matrix. Such direct transfer of observed data is possible if and only if the $\bar{X}\bar{Y}$ cells (probabilities of recall of target to neither cue) in the two data matrices are identical, as they are in Table 2 (.35 in both cases). But even if our basic assumption—that the proportion of items recalled at all (and hence the proportion not recalled at all) is independent of cuing order—is valid, small measurement errors or "noise" effects are to be expected. The critical entries in the data matrices will therefore usually require aligning before they can be combined into a trace matrix.

The most appropriate method of adjusting the data is simple to carry out, but in order to describe it unequivocally it is necessary to adopt a somewhat formal terminology. Consider the two data matrices and the corresponding trace matrix presented in Table 3. These matrices are shown in two forms, one in general symbolic notation, the other containing numerical data from a hypothetical experiment. The symbols X and Y refer to cues X and Y , and subscripts 1 and 2 indicate the order of presentation of the two cues. A plain cue symbol indicates successful recall of the target in the presence of the designated cue, while a symbol with an overbar indicates unsuccessful recall. Valence is in-

¹ The question has been raised whether the derived trace matrix represents the state of the trace at the time of the first probe or at the time of the second probe. There is no satisfactory answer to this question. Since the information in the memory store about any particular item probably changes in the interval between the two probes, our method can only be said to reveal the state of the trace as it "exists" at some time in the interval. The length of the interval between the two probes, of course, can be manipulated in experiment and the accompanying changes in the trace structure observed.

indicated by the symbol V . Thus, for instance, $V(Y_1)$ represents the gross valence of cue Y presented first, $V(\bar{X}_1Y_2)$ represents the reduced valence of cue Y , where cue X was presented first, and so forth. In the following discussion, the symbol V is omitted as redundant. In the trace matrix, the symbols X and Y have no subscripts, since they represent proportions of trace elements independently of cuing order. For instance, $\bar{X}$ represents the proportion of non- X elements in the trace, XY represents the proportion of XY elements, and so forth.

There is, of course, no need to adjust all the entries in the data matrices, but only those (critical) entries that are to be transferred to the trace matrix. The trace matrix entries are derived from the critical entries of the data matrices by using the following equations:

$$\bar{X}\bar{Y} = (\bar{X}_1\bar{Y}_2 + \bar{X}_2\bar{Y}_1)/2 \quad (1)$$

$$= (.38 + .32)/2 = .35,$$

$$X = X_1[(1 - \bar{X}\bar{Y})/(1 - \bar{X}_1\bar{Y}_2)] \quad (2)$$

$$= .43[(1 - .35)/(1 - .38)] = .45,$$

$$Y = Y_1[(1 - \bar{X}\bar{Y})/(1 - \bar{X}_2\bar{Y}_1)] \quad (3)$$

$$= .53[(1 - .35)/(1 - .32)] = .51,$$

$$\bar{X}Y = \bar{X}_1Y_2[(1 - \bar{X}\bar{Y})/(1 - \bar{X}_1\bar{Y}_2)] \quad (4)$$

$$= .19[(1 - .35)/(1 - .38)] = .20,$$

$$X\bar{Y} = X_2\bar{Y}_1[(1 - \bar{X}\bar{Y})/(1 - \bar{X}_2\bar{Y}_1)] \quad (5)$$

$$= .15[(1 - .35)/(1 - .32)] = .14.$$

The equations are illustrated with data from Table 3, with calculations rounded to two decimal places.

The logic underlying Equations 1 to 5 is as follows: The assumption that unsuccessful probing of the trace with a given cue leaves the trace unchanged implies that the order of administering the two cues does not affect the proportion of target items recalled to at least one of the two cues; that is, the commutative law is assumed. The assumption

is stated formally in the following equation:

$$X_1Y_2 + X_1\bar{Y}_2 + \bar{X}_1Y_2 = X_2Y_1 + \bar{X}_2Y_1 + X_2\bar{Y}_1. \quad (6)$$

If Equation 6 is found to hold, then $\bar{X}_1\bar{Y}_2$ will equal $\bar{X}_2\bar{Y}_1$, and the critical values from the data matrices can be entered directly into the trace matrix. If Equation 6 does not hold, then it is assumed that the observed discrepancy between the left- and right-hand sides is attributable to sampling variability and measurement error. The best estimate of the true proportion of items recallable to at least one cue is then given by the mean of the two sides of Equation 6, and each side of the equation is adjusted to this best estimate. It is assumed that measurement and sampling errors are distributed proportionately among the terms of the equation, and hence the adjustment is made by multiplying each term by the ratio of the mean of the two sides to the value of the side of which the term is a part. This adjustment will render the equation true, and consequently produce two adjusted data matrices from which the critical entries can be used to form the trace matrix, exactly as was done in Table 2.

Rather than adjusting the data in the manner just indicated, in practice it is easier to work with the quantities $\bar{X}_1\bar{Y}_2$ and $\bar{X}_2\bar{Y}_1$, which are the complements of the values of the two sides of Equation 6. This is what has been done in Equations 1-5. The underlying logic is exactly as described, but the calculations indicated by Equations 1-5 are simpler.

Actually only three appropriately selected formulas given by Equations 1-5 need be applied; the other values can be obtained by simple subtraction, in keeping with the properties of 2×2 tables. In the illustrative data in Table 2, as well as in most actual experiments, the magnitude of adjustment resulting from the method just described is not very large, thus supporting our basic assumption and suggesting that the violation of the equality in Equation 6 is attributable to noise effects.

TABLE 3
DATA AND TRACE MATRICES

DATA AND TRACE MATRICES

Data matrices							Trace matrix				
First cue X	Second cue Y			First cue Y	Second cue X			Cue X	Cue Y		
	+	-	Total		+	-	Total		+	-	Total
Symbolic notation											
+	$V(X_1Y_2)$	$V(X_1\bar{Y}_2)$	$V(X_1)$	+	$V(X_2Y_1)$	$X(\bar{X}_2Y_1)$	$V(Y_1)$	+	XY	$X\bar{Y}$	X
-	$V(\bar{X}_1Y_2)$	$V(\bar{X}_1\bar{Y}_2)$	$V(\bar{X}_1)$	-	$V(X_2\bar{Y}_1)$	$V(\bar{X}_2\bar{Y}_1)$	$V(\bar{Y}_1)$	-	$\bar{X}Y$	$\bar{X}\bar{Y}$	$\bar{X}$
Total	$V(Y_2)$	$V(\bar{Y}_2)$		Total	$V(X_2)$	$V(\bar{X}_2)$		Total	Y	$\bar{Y}$	
Illustrative data											
+	.36	.07	.43	+	.32	.21	.53	+	.31	.14	.45
-	.19	.38	.57	-	.15	.32	.47	-	.20	.35	.55
Total	.55	.45		Total	.47	.53		Total	.51	.49	

EFFECTS OF ENCODING CONDITION AND RETENTION INTERVAL ON TRACE STRUCTURE: A DEMONSTRATION EXPERIMENT

We are now ready to explore the structure of traces created in experimental situations. We will describe a demonstration experiment in which encoding conditions and the length of retention interval were manipulated as independent variables and then analyze the cued recall data by the reduction method.

Method

Subjects were presented with six different lists of 16 to-be-remembered words each. Each to-be-remembered word was typed in capital letters and accompanied by a context word (input cue) typed in lower case letters. One half of the context words were associatively related to the to-be-remembered words, while the other half were words that rhymed with the to-be-remembered words. For instance, the associative (A) context for the to-be-remembered word CHAIR may have been *desk*, and the rhyming (R) context (presented with the same to-be-remembered word to another subject) may have been *pair*. The two types of context words defined the two encoding conditions of the experiment, and they appeared in the list in an essentially random order. Each list was presented visually just once, at the rate of 2 seconds per word pair. A short digit task was interpolated between the presentation and testing of each list.

Subjects were asked to study each word pair and to prepare for a cued-recall test, with the context words serving as cues for the to-be-recalled target

words. In fact, to encourage subjects to continue to pay attention to the context words throughout the six lists, the target words were indeed tested with their respective input cues after the critical testing phase of each list; we will ignore the data from these input-cue recall tests, since they are of no relevance to our present concern.

The critical data were obtained from the extralist-cue tests interpolated between the presentation of each list and its test with the input cues. In these critical recall tests, each target word was probed twice, once with an associative extralist cue (e.g., *table*), and once with a rhyming extralist cue (e.g., *hare*). Subjects wrote their responses in a test booklet consisting of a number of pages. Different forms of the test booklet were used to manipulate relevant and control irrelevant variables. The first page of the booklet was used for recording the subject's name, the list number, and the digit string. Subsequent pages comprised the cued-recall test. With the exception of the last page of the booklet, which contained the input cues, each test page contained four rhyming cues and four associative cues. Each of the 16 targets was tested on two pages of the booklet, once with an associative extralist cue and once with a rhyming extralist cue.

Of the 16 targets, 4 were tested on the first two pages, 4 were tested on the first and third, 4 on the second and fourth, and 4 on the third and fourth pages. This manipulation of the position of retrieval cues defined the length of the retention interval, the second main independent variable in the experiment. Of each of these subsets of four targets, two had been encoded with respect to an associative context word and two with respect to a rhyming context word. One word in each of these pairs of targets was first probed with a rhyming cue and second with an associative cue, while for the other word the order of cues was reversed.

TABLE 4
TOTAL WORDS RECALLED TO ASSOCIATIVE (A) AND RHYMING (R) CUES AS A FUNCTION OF ENCODING
FORMAT, CUING ORDER, AND PRESENCE OR ABSENCE OF FREE-RECALL TEST

Encoding format	Order of cues	Free-recall test	Both A and R (AR)	Only R (AR)	Only A (AR)	Neither A nor R (AR)
Associated word	A, then R	Yes	231	113	157	267
		No	242	96	159	271
	R, then A	Yes	270	62	172	264
		No	278	67	142	281
Rhyming word	A, then R	Yes	237	136	97	298
		No	232	126	97	313
	R, then A	Yes	234	104	121	309
		No	238	94	126	310

Note. The cell entries for each of the eight conditions sum to 768, for a total of 6144 subject-targets.

Thus for a given subject, each target word within a particular list represented a unique set of experimental conditions. The target words and conditions were balanced within each group of 16 subjects.

The nature of the relation of each cue to its intended target was explicitly specified in the test booklet: For instance, "table (associated with) _____," or "hare (rhymes with) _____." The relation between the two cues of a given target, however, was never specified. Thus, when the subject saw the cue "hare (rhymes with) _____" on page 2 of the booklet, he was not told that the target word sought here was, for instance, the same as the one that belonged to the cue "table (associated with) _____" on page 1.

An additional experimental variable concerned the presence or absence of a free-recall test before the extralist-cue recall test just described. A page containing free-recall instructions was interpolated between the first page and the extralist-cue test pages for half of the booklets of each subject.

A total of 64 Yale undergraduates of both sexes served as subjects. Since each subject was tested with six lists, and each list contained 16 target words, and each target was probed with two extralist cues, the critical data are based on 6,144 subject-words and 12,288 observations.

In summary, to-be-remembered words presented as members of experimental lists were encoded in the context of either associative or rhyming words. Each target word was subsequently probed with two extralist cues, an associative word and a rhyming word. The positions of the cues on the successive pages of the test booklet were systematically manipulated, and so was the presence or absence of an interpolated free-recall test. The data were treated by the reduction method.

Results

The results of main interest concern the effects of encoding conditions and retention intervals on memory traces obtained with

the reduction method. As a general background to these results, Table 4 presents a summary of the total number of words recalled in the eight experimental conditions defined by the orthogonal combinations of the encoding format, the ordering of the associative (A) and rhyming (R) cues, and the presence or absence of the free-recall test; the data are pooled over retention intervals. The data for the free-recall conditions include all words, regardless of whether or not they were produced by the subject in the free-recall test.

Three points are worth noting. The first is that the free-recall test made little difference to the pattern of cuing results. For each combination of encoding format and order of cues, the data are practically indistinguishable for lists tested and not tested by free recall first. Thus, an initial free-recall test does not, or at least in this experiment did not, change the effectiveness of either the associative or rhyming cues.

A second point of interest in Table 4 concerns the effect of encoding conditions on the gross valences of the two kinds of cues. Encoding conditions affected the valence of associative cues, but not the valence of rhyming cues. The total number of words recalled to associative cues was 1,651 under associative encoding conditions and 1,382 under rhyming encoding conditions; the corresponding totals for the rhyming cues were 1,359 and 1,401. These data are very similar to those recently reported by Nelson, Wheeler, Borden, and Brooks (1974).

Finally, Table 4 shows that the violation of Equation 6 was negligible. Within each of the experimental conditions depicted in Table 4, the frequencies of responses in the $\bar{A}\bar{R}$ category (words recalled to neither A nor R cues) were quite similar for the two cuing orders.

For the analysis of principal interest, the data shown in Table 4 were entered into the data matrices, adjusted slightly in order to satisfy Equation 6, and then combined into trace matrices, one for each of eight experimental conditions. Table 5 lists for each of these conditions the four components of the trace, in the form of proportions of the relevant trace elements. The first four experimental conditions in Table 5 are defined in terms of the orthogonal combination of the free-recall and encoding format conditions. The trace components for these conditions were calculated on the basis of the data listed in Table 4. The other four conditions concern only those subject-lists for which the free-recall test was given; they are given by classifying the target items according to whether they had been produced in the free-recall test (some 32% had in fact been produced) on the one hand, and according to their encoding format on the other.

The data in Table 5 (a) confirm the implications in Table 4, both of the lack of any marked effect of the free-recall test and of the effects of the encoding format on the gross valence of the associative but not the rhyming cues, and (b) show that the memory traces of "more difficult" words, those not recalled in the free-recall test, are comparatively weak in the AR component of the trace.

Table 6 is concerned only with target words that subjects failed to produce in the free-recall test, and shows forgetting in the form of changes in trace structure during the course of the retention interval. The early testing refers to the targets cued on pages 1 and 2 of the test booklet, late testing to targets cued on pages 3 and 4.

Forgetting is seen to be reflected primarily in the reduction of the AR elements. Reduced valences of the two cues do not change very much during the testing period, and

TABLE 5
TRACE MATRICES FOR TARGET WORDS IN VARIOUS
EXPERIMENTAL CONDITIONS

Condition	Encoding format	Trace elements			
		AR	AR	$\bar{A}\bar{R}$	$\bar{A}\bar{R}$
No free-recall test given	Associative	.19	.33	.12	.36
	Rhyming	.16	.27	.16	.41
Free-recall test given	Associative	.22	.28	.15	.35
	Rhyming	.16	.26	.18	.40
All words	Associative	.25	.55	.10	.10
	Rhyming	.15	.58	.16	.11
Recalled words	Associative	.21	.15	.17	.47
	Rhyming	.17	.11	.18	.54
Unrecalled words	Associative	.21	.15	.17	.47
	Rhyming	.17	.11	.18	.54

Note. A refers to associative cues, R to rhyming cues.

indeed the reduced valence of the rhyming cues (the $\bar{A}\bar{R}$ component of the trace matrix) seems to have increased from the early to the late tests. The disintegration of the trace revealed by the decrease in the proportion of AR could reflect a transformation of AR elements into $\bar{A}\bar{R}$ and $\bar{A}\bar{R}$ elements at the same time that $\bar{A}\bar{R}$ and $\bar{A}\bar{R}$ elements are transformed into $\bar{A}\bar{R}$ elements, but for the time being this is only conjecture.²

EVALUATION OF THEORY

The theory and method for describing memory traces that we have outlined here, with gross, common, and reduced valences of two (or more) retrieval cues used to specify trace elements and hence trace structure, is really a rather simple-minded one. Indeed its simplicity might be regarded by some as its major weakness. The theory does not predict any phenomena, nor does it explain any, at least in the sense in which

² The fact that the AR elements of the trace are reduced in the course of forgetting and the $\bar{A}\bar{R}$ or $\bar{A}\bar{R}$ elements are not suggests that it is not useful to think of the items contributing data to the AR cell in the trace matrix as "easy" items or as having been recalled mostly by "good" subjects. Were such identification made, the problem would arise of how to explain the forgetting of "easy" but not "difficult" items, or the forgetting by "good" but not "poor" subjects. Any attempt to identify different cells in the trace matrix with items differing in difficulty or with subjects differing in ability also founders on the fact, inherent in the delayed test data in Table 6, that the conditional probability of recall to cue X, given failure of recall to the other cue, Y, is *higher* than the simple probability of recall to cue X uncontaminated by any earlier recall attempt.

TABLE 6

PROPORTIONS OF TRACE ELEMENTS DEFINED BY
ASSOCIATIVE (A) AND RHYMING (R) CUES FOR
WORDS NOT RECALLED IN FREE-RECALL
TEST

Condition	Elements			
	A $\bar{R}$	A R	$\bar{A}R$	$\bar{A}\bar{R}$
Associative encoding				
Early test	.22	.26	.14	.38
Late test	.19	.07	.24	.50
Rhyming encoding				
Early test	.14	.20	.18	.48
Late test	.18	.02	.23	.57

many people think of explanation. It represents little more than an elaboration of the definition of the memory trace in terms of the relation between the conditions and product of retrieval.

On the other hand, the theory enjoys several advantages over others. First, perhaps because of its simplicity, it can be used to define and describe memory traces of all kinds of events, using all kinds of retrieval cues. We have been concerned here with names of list words as the to-be-remembered materials, but there are no reasons why other aspects of list words, larger units of verbal materials, or other kinds of materials altogether could not be subjected to the same logic and the same experimental treatment. For instance, the copy cues (cues nominally identical with target items) used in a recognition test or the "invisible" cues assumed to guide retrieval in free recall could be treated in the same way as the A and R cues in our demonstration experiment. Second, since we define the trace in terms of query-output relations, its description need not emerge in terms of selected features, or as an impoverished copy, of the input item. Third, the trace matrix is a formal expression of the fact that the recall of an event varies with retrieval conditions. Hence there is no logical paradox between a fixed effect of a given experience and its variable memorial manifestations. Fourth, the trace matrix tells us not only about the relative efficiency of different types of retrieval cues but also about how they are related to one

another. Finally, as we have shown in the demonstration experiment, it is a simple matter to find out how encoding conditions affect the trace of one and the same item. We hold the input items nominally constant, vary the conditions under which they are encoded, and then use the reduction method to arrive at the description of the trace.

Our theory as presented here is by no means complete. It will be revised and modified as we learn more about it and as we meet new problems in exploring its utility. In this final section of the article, we will consider some of the more obvious criticisms of the theory.

Trace Structure or Cue Effectiveness?

One of the questions that needs to be considered concerns the identification of various kinds of retrieval cue valences with properties of the memory trace and more specifically with the identification of our trace matrix with trace structure. Why do we refer to a pattern of cue valences as describing the structure of the memory trace, rather than simply discuss the data as they "really are," namely, as measures of cue effectiveness?

We invoke the concept of the memory trace because it helps to integrate, organize, and make sense out of what otherwise would be an unconnected list of experimental procedures and observations. The output from the memory system depends on previous inputs: The system can produce a certain output after but not before a certain input. This means that the system after the input must be somehow different from the system before the input. We can, like many others before us, think of this difference as the memory trace, and this primitive idea helps make sense of the changes in the output from the system. But if we wish to go beyond the mere assertion that an input into the system creates a new memory trace, we have to be more precise about how the trace is to be conceptualized, defined, and described. The theory outlined here attempts to do just this. It provides (a) an internally consistent set of statements about what the memory trace is and what its properties are, and (b) a set of objectively specified

rules by which experimental observations are translated into theoretical terms.

The theory neither demands nor assumes any kind of "real" existence of some "thing" stored in the memory system. Memory trace, as defined by the theory and as described by the reduction method, represents a hypothetical construct that serves to organize a potentially "large list of relations between different questions directed at the system and the output from the system" (Tulving & Bower, 1974, p. 294). When these questions take the form of specific retrieval cues, the trace becomes the *referent* of the cues. We argued earlier that it makes little sense to talk about the effectiveness of a particular cue or the relation between two particular cues. Whenever we state the valence of a cue it is necessary to stipulate its referent, and for us the referent is the trace of a certain event. Thus we talk of valence of cue *X* with respect to trace *T*.

The concept of trace structure, and its description in terms of two or more cues and their informational overlap, represents a straightforward elaboration of these ideas. The measurement and specification of just the gross valences of individual cues provides only partial information about the changes that have taken place in the system as a consequence of a particular input. Probing the memory store with two cues aimed at one and the same target event, and observing the gross, common, and reduced valences of these cues, provides a richer picture of these changes. The similarity or relatedness of any two cues, *X* and *Y*, thus need not be determined by experimenters' intuitions; they can be measured, with respect to a specified trace *T*, by the reduction method. The magnitude of the common valence of two cues in the trace matrix, relative to the reduced valences, provides an objective measure of the extent to which the two cues contain the same retrieval information.

It is for these reasons, then, that we prefer to discuss the memory trace and its structure, rather than simply describe the observed relations between inputs into and outputs from the memory system.

Memory Trace and Memory Organization

Memory trace as defined and described by our theory always represents input events of a particular class. For instance, when input events consist of the presentation of individual words to many people, the resulting trace structure will refer to the population from which the subject-words were sampled, under specified conditions of presentation and recall. It would be a fairly straightforward matter to use just one item and many subjects, or many items and just one subject, and to describe the trace of classes of items defined in this manner. But it is not possible to specify the structure of the trace of a particular event in some particular subject, just as it is not possible to specify the state of one particular atomic nucleus at some particular instant.

Another issue concerns the fact that traces may be stored in an organized fashion (e.g., Mandler, 1967). A critic might wonder whether available evidence for organization in memory invalidates our method, which entails probing for individual items. There is, in fact, no conflict here. The organization of items in memory will be reflected in the trace matrices of these items. A corollary of this fact is that our method provides an approach to the study of organization in memory. For instance, our demonstration experiment showed different trace structures according to whether the target items had been presented in the context of associative or rhyming words (Tables 5 and 6). Further, our method readily lends itself to the investigation of the effects of intralist context effects on trace structure in a free-recall situation. Thus the organization of a given item, or set of items, can be manipulated—for instance, by adding conceptually related items to the list—and the consequent changes in traces described by the reduction method.

Cue Effects Through Mediation?

We have already noted that our theory was developed in the context of the encoding specificity principle: A cue is effective to the extent that its informational contents match the informational contents of the representa-

tion of the experienced episode. It is this principle that permits the specification of trace elements in terms of retrieval cues. But the assumption of a direct match between cue and whatever it is that has been encoded from an earlier event may be questioned.

It could be argued that the effect of a given cue could be mediated through some other cue, and therefore the valences of cues of a designated class provide evidence of questionable relevance to the informational contents of the stored representation of the to-be-retrieved event. The point can be illustrated with our standard example. Assume that the cue "hare (rhymes with) _____" effects recall of the target word CHAIR. In describing the reduction method and its theoretical underpinning, we have assumed that the cue hare is effective because its informational content—its interpretation by the system in semantic memory—somehow appropriately matches the informational content of CHAIR in episodic memory. It is possible to argue, however, that what may have happened is that the cue hare was used in the system to retrieve from semantic memory a number of rhyming words, including the word "chair." Following the implicit retrieval of "chair," its memory location could have been examined for evidence of a particular list tag (or the presence of certain contextual information) and the target word CHAIR recalled because an appropriate tag was identified. Alternatively, the system could have extracted relevant semantic information out of the implicitly retrieved "chair" and matched it with corresponding semantic information about CHAIR in episodic memory. In either case, the effect of the rhyming cue would have been mediated by semantic information, and hence the observed retrieval may tell us nothing about the overlap between the rhyming cue's own unique informational contents and the informational contents of the stored list item.

The reduction method makes it possible to evaluate the mediation hypothesis. The results of our demonstration experiment strongly suggest that, at least in this experiment, the degree of mediation was at most rather small. If a rhyming cue always ef-

fects retrieval of the stored information about the target item by first eliciting a word with semantic contents matching those of the target word, as illustrated in our example above, then the valence of the rhyming cue with respect to the trace reduced by the semantic (associative) cue should be zero. In fact, the proportions of $\bar{A}R$ and $A\bar{R}$ elements in the trace matrices were quite large. A simple mediation hypothesis also founders on the forgetting data shown in Table 6, since it could not explain why the common valence of the A and R cues (AR) drastically decreased over time, while the reduced valences ($\bar{A}R$ and $A\bar{R}$) either stayed the same or even increased.

There is a more fundamental retort to the potential criticism that the effect of a nominal cue may arise through the mediation of some other implicitly generated cue: Whether or not mediation occurs is of little relevance to the logic of our theory. The theory is not inconsistent with the notion that the mechanism of retrieval of stimulus information entails mediation. In fact, the theoretical formulation of the retrieval mechanism, in whatever form, would find relatively few constraints in the theory of trace structure that has been outlined here.

REFERENCES

- Anderson, J. R., & Bower, G. H. *Human associative memory*. Washington, D.C.: V. H. Winston, 1973.
- Anisfeld, M., & Knapp, M. Association, synonymy, and directionality in false recognition. *Journal of Experimental Psychology*, 1968, 77, 171-179.
- Bach, M. J., & Underwood, B. J. Developmental changes in memory attributes. *Journal of Educational Psychology*, 1970, 61, 292-296.
- Barclay, J. R., Bransford, J. D., Franks, J. J., McCarrell, N. S., & Nitsch, K. Comprehension and semantic flexibility. *Journal of Verbal Learning and Verbal Behavior*, 1974, 13, 471-481.
- Bower, G. H. A multicomponent theory of the memory trace. In K. W. Spence & J. T. Spence (Eds.), *The psychology of learning and motivation: Advances in research and theory* (Vol. 1). New York: Academic Press, 1967.
- Elias, C. S., & Perfetti, C. A. Encoding task and recognition memory: The importance of semantic encoding. *Journal of Experimental Psychology*, 1973, 99, 151-156.
- Gardiner, J. M., Craik, F. I. M., & Birtwistle, J. Retrieval cues and release from proactive inhibi-

- tion. *Journal of Verbal Learning and Verbal Behavior*, 1972, 11, 778-783.
- Ghatala, E. S., & Hurlbut, N. L. Effectiveness of acoustic and conceptual retrieval cues in memory for words at two grade levels. *Journal of Educational Psychology*, 1973, 64, 261-266.
- Kintsch, W. *The representation of meaning in memory*. Hillsdale, N.J.: Erlbaum Associates, 1974.
- Kosslyn, S. M., & Bower, G. H. The role of imagery in sentence memory: A developmental study. *Child Development*, 1974, 102, 579-584.
- Light, L. L. Homonyms and synonyms as retrieval cues. *Journal of Experimental Psychology*, 1972, 96, 255-262.
- Mandler, G. Organization and memory. In K. W. Spence & J. T. Spence (Eds.), *The psychology of learning and motivation: Advances in research and theory* (Vol. 1). New York: Academic Press, 1967.
- Morton, J. Interaction of information in word recognition. *Psychological Review*, 1969, 76, 165-178.
- Nelson, D. L., & Brooks, D. H. Relative effectiveness of rhymes and synonyms as retrieval cues. *Journal of Experimental Psychology*, 1974, 102, 503-507.
- Nelson, D. L., Wheeler, J. W., Jr., Borden, R. C., & Brooks, D. M. Levels of processing and cuing: Sensory versus meaning features. *Journal of Experimental Psychology*, 1974, 103, 971-977.
- Rumelhart, D. E., Lindsay, P. H., & Norman, D. A. A process model of long-term memory. In E. Tulving & W. Donaldson (Eds.), *Organization of memory*. New York: Academic Press, 1972.
- Tulving, E., & Bower, G. H. The logic of memory representations. In G. H. Bower (Ed.), *The psychology of learning and motivation: Advances in research and theory* (Vol. 8). New York: Academic Press, 1974.
- Tulving, E., & Thomson, D. M. Encoding specificity and retrieval processes in episodic memory. *Psychological Review*, 1973, 80, 352-373.
- Underwood, B. J. False recognition produced by implicit verbal responses. *Journal of Experimental Psychology*, 1965, 70, 122-129.
- Underwood, B. J. Attributes of memory. *Psychological Review*, 1969, 76, 559-573.
- Underwood, B. J. Are we overloading memory? In A. W. Melton & E. Martin (Eds.), *Coding processes in human memory*. Washington, D.C.: V. H. Winston, 1972.
- Wickens, D. D. Encoding categories of words: An empirical approach to meaning. *Psychological Review*, 1970, 77, 1-15.
- Wickens, D. D. Characteristics of word encoding. In A. W. Melton & E. Martin (Eds.), *Coding processes in human memory*. Washington, D.C.: V. H. Winston, 1972.

(Received September 10, 1974)

A Theory of Attention: Variations in the Associability of Stimuli with Reinforcement

N. J. Mackintosh
University of Sussex, Brighton, England

According to theories of selective attention, learning about a stimulus depends on attending to that stimulus; this is represented in two-stage models by saying that subjects switch in analyzers as well as learning stimulus-response associations. This assumption, however, is equally well represented in a formal model by the incorporation of a stimulus-specific learning-rate parameter, α , into the equations describing changes in the associative strength of stimuli. Theories of selective attention have also assumed (a) that subjects learn to attend to and ignore relevant and irrelevant stimuli (i.e., that α may increase or decrease depending on the correlation of a stimulus with reinforcement) and (b) that there is an inverse relationship between the probabilities of attending to different stimuli (i.e., that an increase in α to one stimulus is accompanied by a decrease in α to others). The first assumption is used to explain the phenomena of acquired distinctiveness and dimensional transfer, the second those of overshadowing and blocking. Although the first assumption is justified by the data, the second is not: Overshadowing and blocking are better explained by the choice of an appropriate rule for changing α , such that α decreases to stimuli that signal no change from the probability of reinforcement predicted by other stimuli.

In experiments on classical conditioning, the arrangement of a contingency between a stimulus and a reinforcer typically establishes that stimulus as an effective conditioned stimulus (CS), which reliably elicits a conditioned response (CR). In experiments on instrumental learning, the reinforcement of a particular response only in the presence of a specific stimulus typically establishes that stimulus as an effective discriminative stimulus, which reliably controls the occurrence of that

response. Whatever differences there may be between these two procedures, it is not unreasonable to expect some similarity in the processes whereby the CS or discriminative stimulus are established as effective signals. The two stimuli may signal different events—the simple delivery of a reinforcer in one case, and the delivery of a reinforcer contingent on the occurrence of a response in the other—but one may hope that similar laws govern the acquisition of these signaling properties. My concern in this article, at any rate, is with the general question of how stimuli are established as signals or acquire associative strength, and no distinction will be made between classical and instrumental experiments.

Virtually all theories of associative learning have taken as their starting point the assumption that increments in the associative strength of a stimulus depend upon some degree of temporal contiguity between occurrences of that stimulus and of the event it signals. By those who wish to express these matters more precisely, it may be assumed that the magnitude

The research reported in this article was supported by grants from the National Research Council of Canada and by the United Kingdom Science Research Council. I am indebted to numerous people who have read earlier versions of the paper or patiently listened to me expound some of the ideas it contains. Although I have no doubt assimilated ideas from many of them, I am particularly conscious of having incorporated points suggested by A. Baker, A. Dickinson, G. Hall, E. Macphail, D. Medin, R. A. Rescorla, and N. S. Sutherland. I am very grateful to them all.

Inquiries concerning this article should be sent to N. J. Mackintosh, Laboratory of Experimental Psychology, University of Sussex, Brighton BN1 9QG, England.

of increments to the associative strength of a particular stimulus, A , is some function of the current value of A 's associative strength, the simplest version of this assumption being that such increments are a linearly decreasing function of that value. This assumption is expressed formally by the application of linear operators to changes in associative strength:

$$\Delta V_A = \theta(\lambda - V_A), \quad (1)$$

where V_A = current associative strength of A ; θ is a learning-rate parameter, the value of which is determined by the reinforcer used within the limits $0 < \theta < 1$; and λ = limit of V , also dependent on the reinforcer used.

This may be a starting point, but it is clearly no more than that. For example, it does not provide any account of one of the most obvious features of both classical and instrumental conditioning, namely, the observation that increments in associative strength depend on the nature of the stimulus. Pavlov (1927) noted that a more intense stimulus was a more effective CS than a less intense one, and that in the case of salivary conditioning with dogs, stimuli from some modalities, such as tactile and auditory stimuli, were more effective CSs than stimuli from other modalities, such as visual or thermal stimuli. Similarly, the wavelength of light to which pigeons are exposed is more likely to gain control over their food-reinforced instrumental responses than is the frequency of a tone sounded from a speaker (Foree & LoLordo, 1973). The simplest way of representing these and similar observations is to assume that the rate of conditioning to a particular stimulus is not just determined by some general learning-rate parameter, such as θ , but is also affected by a stimulus-specific parameter, which may be referred to as α . Thus,

$$\Delta V_A = \alpha_A \theta (\lambda - V_A), \quad (2)$$

where $0 < \alpha < 1$. It is now possible to represent differences in the rate of conditioning to stimuli differing in intensity or salience as differences in the value of α associated with that stimulus. If condi-

tioning proceeds faster with stimulus A than with stimulus B , or if subjects, given conditioning trials with a compound stimulus consisting of $A + B$, respond strongly to A but weakly to B , these facts may be represented by assigning a high value to α_A , and a low value to α_B .

To this point, the elementary ideas being advanced do not represent any significant departure from the central assumptions of traditional theories of learning, at least of theories such as those advanced by Hull (1943), Spence (1936, 1956), or Estes (1950, 1959). At least two recently proposed models in this tradition, for example, have included a stimulus-specific learning-rate parameter such as α (Kendler, 1971; Rescorla & Wagner, 1972).

THEORIES OF SELECTIVE ATTENTION

The framework expressed by Equation 2 is, in fact, sufficiently general to include another group of theories, superficially quite distinct. Theories of selective attention, of which several versions have been advanced in recent years (for example, Lovejoy, 1968; Sutherland, 1964; Sutherland & Mackintosh, 1971; Trabasso & Bower, 1968; Zeaman & House, 1963), appear to incorporate some quite radically new assumptions about the processes involved in even simple cases of instrumental learning. Nevertheless, I shall argue that their two central assumptions can easily be represented as variants on Equation 2.

Theories of selective attention have all been formalized as "two-stage" or "chaining" models, where the subject is envisaged as first observing or attending to a set of stimuli on each trial, and then making an overt response determined only by those stimuli attended to on that trial. In this form, they seem to differ sharply from the "single-stage" type of model characterized by Equation 2. Yet although the function of the observing or attentional stage may be described informally as that of initially exposing the subject to a particular set of discriminative stimuli so that those stimuli and no others are available for association with the outcome of the

trial, formally it is only this second point—whether or not a particular stimulus has its associative strength changed—that is of any substance.¹

This point has been represented in two different ways. In theories such as those of Zeaman and House (1963) and Lovejoy (1968), the subject is assumed to attend to only one set of stimuli on a trial. The probability of attending to a particular stimulus on a particular trial, therefore, determines the probability that the outcome of that trial will change the associative strength of that stimulus. In the theory of Sutherland and Mackintosh (1971), on the other hand, subjects are assumed to switch in several analyzers on a given trial. The strength of a particular analyzer, therefore, determines the amount by which the associative strength of a particular stimulus is changed. Over a series of trials, of course, these two formalizations amount to very much the same thing. More important, however, insofar as the strength of switching in a particular analyzer or the probability of attending to a particular set of stimuli simply determines which stimuli change their associative strengths (and by how much), this characteristic of two-stage models is equally well represented by a single-stage model that incorporates a learning rate parameter, α , as exemplified in Equation 2. Although theories of selective attention have, indeed, incorporated further assumptions, these assumptions are also easily represented in a single-stage model as determinants of the value of α .

The first of these assumptions, common to all theories of selective attention, is that α is not a fixed consequence of such physical characteristics of a stimulus as its intensity or modality, but rather that α may vary with the subject's experience. The experience in question may have occurred before the start of an experiment,

but in the more interesting case, it is a consequence of the correlation between the stimuli and the reinforcement to which the subject is exposed during the course of an experiment. In the language normally used by such theorists, the probability of attending to relevant stimuli typically increases, while the probability of attending to irrelevant stimuli typically decreases. In other words, if variations in stimulus A are correlated with changes in reinforcement, α_A may increase, while if variations in stimulus B are not correlated with changes in reinforcement, α_B may decrease.

The second major assumption common to all of these theories is that the probability of attending to one set of stimuli is inversely related to the probability of attending to others. This is a necessary consequence of the "one-look" assumption incorporated into the models of Zeaman and House (1963) and Lovejoy (1968). Since subjects are assumed to attend to only one set of stimuli on a trial, it necessarily follows that as the probability of attending to one set increases, so the probability of attending to others must decrease. Moreover, even Sutherland and Mackintosh (1971), although allowing that learning can occur about all available stimuli on a single trial, assume that the strengths of all analyzers sum to 1.0. In general, therefore, theories of selective attention may be said to have assumed an inverse relationship between the values of α specific to the population of stimuli sampled in any given experimental situation. If that situation can be conceptualized as containing only three sets of stimuli, A, B, and X, an increase in α_A will be accompanied by corresponding decreases in either or both of α_B and α_X . This assumption has been called the inverse hypothesis (Thomas, 1970).

Although these two assumptions (that the value of α rather than being fixed may change with experience, and that the values of α for different stimuli are inversely related to each other) have both been incorporated into essentially all theories of selective attention, and although

¹ This is of course an oversimplification, for the direction of the subject's attention is also assumed to affect which stimuli will actually determine performance at any given moment, but the question of performance may for the moment be deferred.

both have also been rejected by the majority of other theories of associative learning, it is important to see that they are logically independent. There is, in fact, no reason why the possibility that subjects might learn to attend to relevant stimuli or ignore irrelevant stimuli should carry with it the implication that such attention is *selective* in the sense required by the inverse hypothesis. One can conceive of a theory of associative learning, which while permitting α for various stimuli to change in accordance with the schedule of reinforcement associated with those stimuli, does not insist that a change in α to one set of stimuli is necessarily accompanied by an equal and opposite change in α to other sets of stimuli. Since a theory that accepts one of the assumptions common to most theories of attention without the other is thus clearly a logical possibility, it is important that the pre-suppositions and experimental evidence relevant to the two assumptions be examined independently.

Changes in Attention with Experience

Traces of the assumption that the direction of attention may change can be found in both Lashley's (1929) and Krechevsky's (1932) early writings on discrimination learning. According to Krechevsky, subjects entered the experimental situation with a particular set of hypotheses about which stimuli might be relevant and which irrelevant, but in the face of inconsistent reinforcement would discard unsuccessful hypotheses and try out new ones. Although hypotheses were thus said to be "docile," the actual rules governing changes from one hypothesis to another were not specified, and the implication was that they were better characterized as consequences of a sudden new insight, or of the reorganization of the subject's perceptual field, rather than being due to any readily understood process of associative learning. In this context, it was Lawrence (1949, 1950) who first suggested that the "acquired distinctiveness of cues" might be a gradual learning process, amenable to analysis in traditional terms.

It was also Lawrence, of course, who first provided relatively convincing evidence that changes in α might occur. His experiments on simultaneous and successive discriminations (Lawrence, 1949, 1950) and on transfer along a continuum (Lawrence, 1952) showed transfer from one discrimination problem to another, which seemed to imply increases in the value of α for relevant stimuli and/or decreases in the value of α for irrelevant stimuli. This is not the place for a detailed review of these studies, but although other accounts of these transfer effects have been offered, there is reason to question their sufficiency (Sutherland & Mackintosh, 1971). Similarly, although experiments on discrimination reversal undoubtedly involve a number of complex processes, it can be argued that one reason why overtraining sometimes facilitates reversal learning, and why exposure to a series of reversals of a single discrimination sometimes leads to eventually very rapid reversal, is that both overtraining and serial reversal training increase the value of α for the relevant discriminative stimuli (Mackintosh, 1969; Mackintosh & Holgate, 1969). Finally, studies of intradimensional and extradimensional shifts have shown that the magnitude of transfer from one discrimination problem to another depends on whether the stimulus dimensions that were relevant and irrelevant in the first problem retain their original significance in the second problem (Shepp & Eimas, 1964; Shepp & Schrier, 1969). It is very difficult to see how such a finding can be explained without assuming changes in α for relevant and/or irrelevant stimuli, which transfer to other stimuli falling along the same dimension.

It may be accepted that many of these findings imply that α changes during the course of discrimination learning; they do not, however, enable one to specify the direction of such changes. As was implicitly acknowledged in the preceding paragraph, many of these results could as well be a consequence of a decrease in α for irrelevant stimuli, as of an increase in α for relevant stimuli. It is somewhat

surprising, therefore, to find that formal theories of selective attention have concentrated on this second interpretation to the virtual exclusion of the first. It is true that Lawrence (1949) admitted that a decline in the distinctiveness of irrelevant stimuli was probably at least as important as an increase in the distinctiveness of relevant stimuli in generating the transfer effects he observed. But in the formal two-stage models of discrimination learning proposed by Zeaman and House (1963), Lovejoy (1968), and Sutherland and Mackintosh (1971), while differential reinforcement is assumed to increase attention to relevant stimuli, there is no comparable independent mechanism for reducing attention to irrelevant stimuli. Except in the sense that a formerly relevant analyzer will revert to its baseline strength if the stimuli it detects subsequently become irrelevant, the only mechanism for decreasing the strength of irrelevant analyzers is by competition with relevant analyzers. It is, in fact, only the inverse hypothesis that ensures that as attention to relevant stimuli increases, so attention to irrelevant stimuli must decrease.

There are good reasons to question this analysis. A substantial body of evidence suggests that exposure to a set of stimuli uncorrelated with reinforcement will specifically reduce the associability of those stimuli with subsequent changes in reinforcement, and that this is not just a consequence of any increase in attention to another set of stimuli. The simplest demonstration of this effect is provided by experiments on latent inhibition (Lubow, 1973; Lubow & Moore, 1959), in which nonreinforced preexposure to a particular stimulus interferes with subsequent conditioning to that stimulus. Experiments by Rescorla (1971) and Reiss and Wagner (1972) have established that such nonreinforced preexposure does not turn a stimulus into a conditioned inhibitor or signal for nonreinforcement and indeed may significantly interfere with the establishment of subsequent inhibitory conditioning. It is difficult, therefore, to see how such results can be explained except

by postulating changes in a stimulus-specific learning-rate parameter such as α . Nor is it necessary to present a stimulus in the complete absence of reinforcement, as in typical experiments on latent inhibition, in order to retard subsequent conditioning. Indeed, sufficient exposure to uncorrelated presentations of a stimulus and reinforcer may be an even more effective procedure for reducing the associability of that stimulus with that reinforcer (Mackintosh, 1973).

Latent inhibition has also been reported in experiments on free-operant instrumental discrimination learning by Mellgren and Ost (1969) and Halgren (1974). Although a similar finding has not been reported in discrete-trial studies, there is evidence of what would seem to be an analogous effect. In experiments on simultaneous and successive discrimination learning, Waller (1970) and Hall (Note 1) obtained results that suggest that rats and pigeons may learn specifically to ignore irrelevant stimuli that vary from trial to trial without being correlated with reinforcement.

There is, therefore, a considerable amount of evidence, consistent in a general way with the first assumption of theories of attention, that the associability of a stimulus with reinforcement is not an immutable consequence of that stimulus's physical characteristics. Exposure to a set of stimuli correlated with changes in reinforcement appears to increase the associability of those and similar stimuli with future changes in reinforcement. Conversely, exposure to variations in a set of stimuli in the absence of correlated changes in reinforcement appears to decrease the associability of those stimuli with future changes in reinforcement. The fact that such changes in α occur, of course, does not of itself specify the nature of the rules for changing α that a formal model would have to incorporate in order to provide a satisfactory account of the data. But there is already a strong suggestion that it is not enough to assume that attention to irrelevant stimuli decreases simply by virtue of an increase in attention to relevant stimuli. It seems probable that

a symmetrical set of rules will be required that will permit opposite but independent changes in α to relevant and irrelevant stimuli.

The Inverse Hypothesis

The second main assumption held by most theories of selective attention is that stimuli compete for attention, so that an increase in attention to one set of stimuli will necessarily involve a decrease in attention to some other set of stimuli. It should be clear that the experiments reviewed in the preceding section, designed to show that the probability of attending to a particular set of stimuli may change with experience, do not have any bearing on the validity of this second assumption. What then has been the justification for the inverse hypothesis?

One line of argument has been that it is a necessary corollary of the assumption, derivable from information theory, that there are limits to any organism's capacity to process information. Stimuli must be assumed to compete for access to a channel of limited capacity, and the formation of associations between one set of stimuli and reinforcement will necessarily be at the expense of the formation of associations with other stimuli. Plausible as it may seem to insist on limitations on the capacity to process information, however, it may be questioned whether these concepts are of any substantial relevance to the sorts of situations to which theories of selective attention have been applied. Of the numerous differences between experiments on instrumental discrimination learning in animals, and those studies of dichotic listening or reaction times in human subjects on which the concept of a limitation on capacity has been based, perhaps the most important is that in the former situation subjects control their own exposure to the discriminative stimuli; they are under no constraints to respond rapidly, and there is little reason to suppose, therefore, that they are processing information at rates close to the limits of their capacity. Although it is true that in experiments on classical conditioning

the experimenter controls the subject's exposure to the CS, the stimuli used are typically salient, often unlocalized, and may sometimes last up to several minutes. It stretches credulity to suppose that the rat's channel capacity is too limited to permit processing of both a loud tone and a bright light, when the two are switched on simultaneously for the duration of a 3-minute trial in an experiment on conditioned suppression. Whether or not strict simultaneity of stimulus processing is impossible, typical animal experiments have imposed no such requirement.

It is obvious, however, that the major justification for a particular theoretical assumption is the extent to which it permits the derivation of experimental data. What are the data that are typically explained by the inverse hypothesis? They may generally be described as instances of "stimulus selection," data which suggest that the association of one stimulus with reinforcement does not depend simply on the characteristics of that stimulus, on its correlation with reinforcement, or even, apparently, on its past history, but may be affected by the nature of the other stimuli present at the same time, by their correlation with reinforcement, and by their past history. The most important instances of such data are provided by studies of overshadowing and blocking. The term *overshadowing* refers to the finding, first reported by Pavlov (1927), that the presence of an equally relevant, more salient stimulus may decrease or completely prevent conditioning to a less salient stimulus. A dog would associate a weak thermal stimulus with the delivery of food, provided it was presented on its own, but if it was presented only in conjunction with a more intense auditory stimulus, no conditioning would accrue to the weaker stimulus. In general, conditioning to B often proceeds more slowly if B is always presented in conjunction with A than if B is presented alone.

The term *blocking* was used by Kamin (1969) to refer to the finding that prior training on one element, A, of a compound stimulus, AB, might completely

prevent conditioning to the second element, B. If rats received a series of trials on which a compound CS, consisting of a light and a noise, signaled the occurrence of a shock, they would show significant conditioned suppression to both stimuli. But if these compound trials were preceded by a series of trials on which, say, the light alone signaled shock, they would show little or no suppression to the noise.

Both overshadowing and blocking, it need hardly be pointed out, follow quite directly from the inverse hypothesis. An intense stimulus is precisely one that is associated with a high value of α , and its addition to the stimulus compound correlated with reinforcement will necessarily, according to the inverse hypothesis, decrease the value of α associated with any other stimulus. Similarly, prior training with A relevant will increase α_A and hence ensure that when B is added, α_B will be low. In spite of this apparently happy agreement between data and theory, however, one can still ask whether blocking and overshadowing are explained more satisfactorily by a theory of selective attention than by any other theory, and whether they are observed on the occasions required, and for the reasons specified, by such a theory. There is reason to believe that they are not.

Rescorla and Wagner (1972) proposed an alternative explanation of these and other instances of stimulus selection. Kamin (1969) suggested that blocking occurs because on compound trials the subject already expects the reinforcer and only unexpected reinforcers are reinforcing. Rescorla and Wagner have formalized one aspect of this idea: The central assumption of their model is that changes in the associative strength of a stimulus are inversely related, not just to the current associative strength of that stimulus, but to the current strength of the entire stimulus complex of which it forms a part. Formally,

$$\Delta V_A = \theta(\lambda - \bar{V}), \quad (3)$$

where $\bar{V}$ is the sum of the associative

strengths of all stimuli (including A) present at the moment of reinforcement. In an experiment on blocking, a sufficient number of reinforced trials with A alone will ensure that V_A approaches λ ; at this point, reinforced trials with AB will result in little or no conditioning to B, since $(\lambda - \bar{V})$ will be nearly zero.

Rescorla and Wagner's model, although clearly quite distinct from any theory of selective attention, may still, without gross distortion, be said to explain stimulus selection by appeal to an inverse hypothesis. Stimuli are assumed to compete with one another, not for a limited attentional capacity but for a limited amount of associative strength conditionable by a given reinforcer. Other things being equal, the greater the associative strength accruing to one element of a compound CS, the less will be the associative strength of the other. Unlike theories of selective attention, however, Rescorla and Wagner's theory does not postulate what may seem to some implausible limits to an animal's capacity to attend to incoming stimuli; moreover, as we shall see, there are cases where its predictions differ from those derivable from theories of selective attention and cases where it is supported by the evidence. I shall argue, however, that in the final analysis, the data on stimulus selection do not support any version of the inverse hypothesis: Blocking and overshadowing are not the consequence of any simple competition between stimuli for some limited resource.

EXPERIMENTAL ANALYSIS OF STIMULUS SELECTION

Blocking

If A is established as a signal for reinforcement, and subjects are then exposed to AB, signaling the same reinforcer, the available data suggest that the failure of conditioning to B is not due to the fact that A preempts attention but is rather a consequence of the fact that B signals no change in reinforcement, predicting nothing that is not already predicted by the presence of A. When this condition

is not satisfied, blocking may not occur. In Kamin's (1969) experiments on blocking in conditioned suppression, if A is established as a signal for a 1 mA shock, and AB also signals a 1 mA shock, B acquires relatively little associative strength. But if AB signals a stronger shock than that signaled by A alone, or no shock at all, then significant excitatory or inhibitory conditioning accrues to B. In the latter case, Wagner (1971) has shown that if reinforced trials with A alone are followed by nonreinforced trials with AB, then an increase in the number of reinforced trials to A, so far from decreasing α_B and the amount learned about B, actually ensures that B is established as an even stronger conditioned inhibitor.

While it might still be possible in the face of this evidence to save the analysis of blocking suggested by theories of selective attention, for example by assuming that the values of α change when the conditions of reinforcement change, it certainly seems more appropriate to assume that any adequate explanation of blocking must take as its starting point the cardinal observation that blocking depends on the redundancy of the added element: Blocking occurs because AB signals the same reinforcer as that signaled by A alone. This is, of course, precisely the position adopted by Kamin (1969) and Rescorla and Wagner (1972). But there is another way of explaining blocking in terms of redundancy, one which appeals neither to limitations of attentional capacity nor to limitations on the total associative strength conditionable by particular reinforcers. It is possible that the redundancy of B reduces conditioning to B because it reduces α_B .

I have already argued that any adequate theory of associative learning must allow α to change and that this change must include a rule for decreasing α to stimuli uncorrelated with changes in reinforcement. If it were assumed that α_A increases whenever A signals a change from the prevailing or expected conditions of reinforcement, the symmetrical rule for decreasing attention to irrelevant stimuli

would say that α_B decreases whenever B signals no change from the prevailing or expected rate of reinforcement. The expected rate of reinforcement, however, is precisely that which is predicted by other stimuli in the experimental situation. Thus the blocking experiment, in which AB signals no change in reinforcement from that predicted by A alone, is precisely the sort of situation that would result in a rapid decline in α_B .

It is, in fact, possible to show that such a decline in α_B is one of the things that happens when AB signals the same reinforcement previously signaled by A alone. It will be recalled that if animals have learned that A signals a 1 mA shock, the presentation of AB, correlated either with a significant increase in the intensity of shock or with no shock at all, results in significant excitatory or inhibitory conditioning to B. Mackintosh and Turner (1971) found that a small number of trials on which AB signaled the same shock as A alone, interposed between initial conditioning to A and subsequent training with AB, significantly impaired this excitatory or inhibitory conditioning. The implication is that α_B decreased because when B was first introduced it signaled no change in the probability or magnitude of reinforcement.

It is, of course, one thing to say that if AB signals the same reinforcement previously signaled by A alone, α_B will decrease. It is another thing to say that this is the sole or even a major cause of blocking. *This* assertion has a number of further implications. If the blocking of conditioning to B depends on subjects learning to ignore B, for example, no blocking can be expected until they have done so. Where blocking is substantial, it must be assumed that changes in associative strength are slow relative to changes in α . If conditioning proceeded more rapidly, one might expect to see relatively little blocking. But even under normal circumstances, one would expect to see some conditioning to B, since it must take some minimal number of trials (at least one) to reduce α_B to a point

where further changes in the associative strength of B are minimal.

There is, in fact, some evidence consistent with this final implication. Kamin (1969) found it possible to detect some conditioning to B in his standard experimental situation; more important, his data suggested that most if not all of this conditioning was a consequence of the first reinforced AB trial. This observation agrees with some results of an experiment on blocking in instrumental learning by Turner (Note 2). In this experiment, although pretraining on A resulted in some blocking of learning about B during AB trials, blocking was far from complete. Since animals receiving 100 AB trials learned no more about B than those receiving only 10 AB trials, however, the implication is that learning about B was confined to the first few trials after its introduction.

Although these results are consistent with the present argument, they do not bear on the strong prediction derivable from this analysis, that little or no blocking should occur on at least the first reinforced compound trial. A satisfactory test of this prediction requires a situation in which reliable conditioning occurs to the elements of a compound CS after a single reinforced trial. The conditioned suppression of licking in thirsty rats by a CS signaling shock provides one such situation, and we have attempted to see whether under these circumstances pretraining on A does result in significant blocking of conditioning to B when only one reinforced AB trial is given (Mackintosh, Note 3). In several experiments, no evidence of blocking could be found. These results suggest, therefore, that a substantial part of blocking observed in other situations may indeed be a consequence of a rapid decline in attention to the added, redundant element.

Overshadowing

The design of experiments on blocking is such that the added stimulus B predicts no change in the level of reinforcement from that predicted by A alone, and it

was, of course, precisely this feature that made it possible to explain blocking in terms of a reduction in α , rather than in terms of the inverse hypothesis. But in experiments on overshadowing, animals may be trained with AB-signaling reinforcement from the outset. If A has α already been established as a signal for reinforcement, how can its presence detract from conditioning to B during reinforced compound trials, unless some appeal is made to the idea that stimuli compete with each other for association with reinforcement? If a subject simply receives reinforced trials with an AB compound, both A and B should be established as signals for reinforcement, and without invoking the inverse hypothesis it is hard to see why the presence of one should cause any decline in attention to the other.

There are, however, at least two circumstances that would appear to provide a sufficiently close parallel to the conditions of an experiment on blocking to make it possible to apply the principle of learned irrelevance to the case of overshadowing. If A is better correlated with reinforcement than B, or if A is a much more salient stimulus than B and therefore conditions more rapidly, then there is a sense in which B will be a redundant signal of reinforcement. In the former case, A, by virtue of its superior schedule of reinforcement, will acquire associative strength more rapidly than B. In the latter case, even though the schedule of reinforcement associated with A and B is identical, since A conditions more rapidly than B, A will be established as a reliable signal for reinforcement after a number of reinforced compound trials, while B will have acquired little associative strength. At this point, just as in the case where the schedule of reinforcement associated with A is better than that associated with B, the situation will be analogous to that obtaining in experiments on blocking: The presence of B predicts nothing that is not already predicted by A alone, and α_B will decline.

Numerous studies of compound conditioning have established that of two equally salient stimuli, the more valid

stimulus, associated with a better schedule of reinforcement, will overshadow the less valid (e.g., Egger & Miller, 1963; Wagner, 1969; Wagner, Logan, Haberlandt, & Price, 1968) and that of two equally valid stimuli, the more salient will overshadow the less salient (e.g., Kamin, 1969; Miles Jenkins, 1973; Pavlov, 1927). What has not been recognized, however, is that these appear to be the only reliably established instances of overshadowing. Although the more salient element of a compound CS overshadows the less salient, there is no evidence that the less salient will detract from conditioning to the more salient, nor is there much evidence to suggest that if two equally salient stimuli are associated with the same schedule of reinforcement, one will overshadow the other.

If stimuli compete for association with reinforcement in the manner suggested by the inverse hypothesis, one would expect to observe a reciprocal interaction between all elements of a compound CS. If animals receive a series of reinforced trials with AB, then, other things being equal, any conditioning to B must always be at the expense of conditioning to A, and vice versa. The experimental evidence does not support this expectation. Pavlov (1927) reported that overshadowing was a matter of a reduction in conditioning to the weaker of two elements of a compound CS as a consequence of the presence of the stronger. The unspoken implication was that conditioning to the stronger was unaffected. Although Pavlov did not present the data to support this inference, the conclusion is entirely consistent with subsequent experiments.

Kamin (1969) and Mackintosh (1971), in studies of conditioned suppression, reported that a bright light would overshadow auditory stimuli of varying intensities, but neither found any evidence to suggest that the presence of the auditory stimulus detracted from conditioning to the light, nor that the amount of conditioning to the light was an inverse function of the intensity of the auditory stimulus. Schnur (1971), using relatively

intense auditory and visual stimuli in another study of conditioned suppression, found no evidence of overshadowing at all. It is true that in some of these experiments ceiling effects may have obscured any evidence of overshadowing. In order to obtain further evidence on this point, we have studied overshadowing in conditioned suppression, using noise and light as CSs, and varying the intensity of the noise between 50 and 85 db. re 20 $\mu\text{N}/\text{m}^2$. To counteract ceiling effects, subjects were tested over a series of trials in extinction, to the point where no group maintained complete suppression. Although there was some evidence of reciprocal overshadowing at intermediate intensities of the noise, in the extreme cases there was no evidence of any such interaction. The most intense noise significantly overshadowed the light, while the least intense noise was itself significantly overshadowed by the light. In neither instance, however, did the overshadowed element, even though itself acquiring significant associative strength, detract from conditioning to the stronger element.

In experiments on instrumental discrimination learning, significant overshadowing has been observed, provided that the overshadowed cue is either extremely difficult to discriminate (Lovejoy & Russell, 1967) or both relatively difficult and in a different location from the more salient, overshadowing cue (Sutherland & Andelman, 1967). When neither of these conditions is satisfied, overshadowing is not observed (Sutherland & Andelman, 1967; Turner, 1968), and it may be difficult to detect any evidence of an interaction between the stimuli (Sutherland & Holgate, 1966; Warren, Derdzinski, Hirayoshi, & Mumma, 1970). The most extensive study of overshadowing in instrumental discrimination learning is that by Miles and Jenkins (1973). Different groups of pigeons were trained on a successive discrimination with either the presence or absence of a tone, a difference in light intensity, or a combination of both cues serving to distinguish posi-

tive and negative trials. The discriminability of the light cue varied across groups. The most striking feature of their results is that the presence of the tone detracted from control by the light only with the least discriminable value of the light, but for this group the presence of the light did not significantly detract from control by the tone. Conversely, the more discriminable lights did overshadow the tone, but in these groups the presence of the tone did not detract from control by the light. No group, in other words, showed simultaneous overshadowing of both light by tone and tone by light.

It is too early to reach any unequivocal conclusion, but the weight of the evidence does not appear to provide much support for the inverse hypothesis. Overshadowing does not appear to be the general, reciprocal affair required by any theory that assumes stimuli necessarily compete for a limited total amount of associative strength or for access to a limited-capacity channel. To a first approximation, stimuli that are overshadowed appear to be only those with lower salience or inferior correlation with reinforcement, which results in a slower rate of conditioning than that accruing to concurrently presented, overshadowing stimuli. Any interaction between elements of a compound that is confined to this sort of situation may be quite satisfactorily explained by a theory that postulates reductions in α to stimuli that predict no unexpected changes in reinforcement.

SPECIFICATION OF A NEW THEORY OF ATTENTION

If the data traditionally thought to support one of the assumptions adopted by theories of selective attention can be explained in other ways, it may be time to drop that assumption. The theory that emerges from the considerations advanced so far may thus be characterized as follows. Changes in the associative strength of a stimulus do not depend only on its correlation with reinforcement or on the magnitude of that reinforcement; they also depend on the nature of the stimulus.

Stimuli may condition at different rates, and at least one cause of these differences may be represented, as in Equation 2, by differences in a stimulus-specific learning-rate parameter, α . The value of α is initially determined by the physical characteristics of the stimulus and the subject's sensory apparatus, but it may also change with experience. In particular, if stimulus A is correlated with changes in reinforcement, α_A will increase, and if stimulus B is not correlated with changes in reinforcement, α_B will decrease. The critical departure from traditional theories of selective attention is that in a situation where several stimuli, A, B, C, are presented, an increase in α_A does not necessarily or directly cause a decline in α_B or α_C . Changes in α_B and α_C are independent of changes in α_A . They are caused solely by the correlation of B and C with changes in reinforcement and do not depend in any direct way on changes in α_A .

Rules for Changing α

Even if it is assumed that there may be changes in attention to particular stimuli, the assumption that such changes in attention are entirely independent at first sight seems to rule out the possibility of explaining the types of interaction between stimuli that are exemplified by the phenomena of blocking and overshadowing. This is the conclusion accepted by Fisher and Zeaman (1973), who have recently proposed a modification of the original Zeaman and House (1963) model, one of the new features of which is a rejection of the inverse hypothesis. They noted that their revised theory would not predict blocking or overshadowing under normal conditions. As the arguments of the preceding sections perhaps have shown, the validity of this conclusion depends on the nature of the rules proposed for changing attention. Fisher and Zeaman (1973) follow Sutherland and Mackintosh (1971) in assuming that the strength of a particular observing response will increase whenever the outcome of a trial confirms the subject's expectations about the stimulus detected by that observing response;

similarly, disconfirmation of an expectation results in a decrease in the strength of that observing response. Rules such as these will not, of course, predict any decline in the strength of an observing response to the redundant stimulus in an experiment on blocking. If a subject receives a series of reinforcements signaled by AB, the fact that these trials were preceded by a series of reinforced trials signaled by A alone will have no effect on any changes in the probability of attending to B. The rules for changing observing responses make no reference to whether a particular reinforcer is or is not expected on the basis of other simultaneously presented stimuli.

A theory such as Fisher and Zeaman's, therefore, in which observing responses change independently and simply in accordance with the reinforcement schedule associated with each independent set of stimuli, while able to explain data on the acquired distinctiveness of cues, reversal learning, and intradimensional and extradimensional shifts, will not predict such phenomena of stimulus selection as blocking and overshadowing. It is possible, of course, that such phenomena are in no sense attentional in origin. One might argue that the phenomena of acquired distinctiveness and stimulus selection could (and should) be handled by entirely different theoretical constructs. An adequate theory of associative learning would be one that explained acquired distinctiveness by allowing α to change in the manner proposed by Fisher and Zeaman (1973), and that explained stimulus selection in the manner proposed by Kamin (1969) and Rescorla and Wagner (1972).

Although a case could be made for such a hybrid theory, it must be recognized that Rescorla and Wagner's analysis of blocking and overshadowing is a special case of the inverse hypothesis, since increases in the associative strength of one component of a compound CS are assumed to be at the expense of other components. As we saw earlier, therefore, the analysis is vulnerable to some of the arguments that can be advanced against any such

interpretation of stimulus selection. If overshadowing is not the general and reciprocal phenomenon postulated by the inverse hypothesis, and if little or no blocking occurs on the first trial on which a new component is added, this is evidence against Rescorla and Wagner's analysis just as much as against the version of the inverse hypothesis postulated by theories of selective attention. The force of the previous argument was precisely that overshadowing and blocking are consequences of reductions in the value of α associated with overshadowed and blocked stimuli and are therefore to be explained by the postulation of appropriate rules for changes in α .

It remains, then, to propose such rules. Instead of saying, as do Sutherland and Mackintosh (1971) and Fisher and Zeaman (1973), that subjects learn to attend to and ignore stimuli to the extent that those stimuli successfully predict the outcome of a trial, we want to say that it depends on whether the stimuli are *uniquely* successful in their predictions. The intuition that we require to formalize is that α_A should increase if A predicts an otherwise unexpected reinforcer, while α_A should decrease if A signals no change in reinforcement from the level expected on the basis of other events. There are presumably, a number of ways in which this might be done, but possibly the simplest is as follows. The extent to which a reinforcer is predicted by A is represented by the absolute value of the term $|\lambda - V_A|$, where, as usual, λ is the asymptotic associative strength conditionable by that reinforcer, and V_A is the current associative strength of A. If we wish α_A to increase whenever the outcome of a trial is predicted by A better than by all other events on that trial, we could say,

$$\Delta\alpha_A \text{ is positive if } |\lambda - V_A| < |\lambda - V_X|, \quad (4)$$

where V_X is the associative strength of all stimuli other than A present on that trial. Conversely, if we wish α_A to decrease whenever the outcome of trial is predicted by other events at least as well as by A,

we could say,

$$\Delta\alpha_A \text{ is negative if } |\lambda - V_A| \geq |\lambda - V_X|. \quad (5)$$

Stimuli can, of course, be established as signals for nonreinforcement just as for reinforcement. The simple assumption that the value of λ for nonreinforcement is either zero or some negative number will permit appropriate changes in α on nonreinforced trials.²

It is easy to see how such rules will predict the occurrence of overshadowing and blocking under appropriate circumstances; α_B will decrease if B is always presented in conjunction with A and A has previously been established as a signal for the reinforcer in question, or if, by virtue of its greater salience, A is initially associated with a higher value of α and thus acquires associative strength faster than B. If it is further assumed that the size of the change in α is proportional to the discrepancy between $|\lambda - V_B|$ and $|\lambda - V_X|$, then α_B will decrease most rapidly in the blocking situation, where on the first trial on which B is introduced, V_B is near zero and V_X (which includes V_A) is near λ . Conversely, α_A will increase faster if reinforcement is signaled by A alone than by a compound CS that includes other stimuli as well as A. In principle, therefore, some reciprocal overshadowing would be predicted, but not if the stimuli were sufficiently salient to be associated with high values of α at the outset of conditioning. The data on the effects of salience on overshadowing presently available are not sufficient to say whether such an assumption is justified; it might prove necessary to adopt rather different rules in light of further research.

² This should not be construed as unconditional acceptance of Rescorla and Wagner's (1972) assumption that reinforcement and nonreinforcement both result in changes in a single variable, V . For present purposes, one could equally well assume that the omission of an expected reinforcer resulted in an increment in some specific inhibitory process, which grew to some asymptote, λ_i , and that the net associative strength, V , of a stimulus was determined by subtracting this inhibitory process from a separate excitatory process which was incremented on reinforced trials.

Overshadowing will also be predicted if A is a better predictor of reinforcement than B, for example if reinforced trials signaled by AB are alternated with nonreinforced trials signaled by B alone. Under these conditions, the different schedules of reinforcement associated with the two stimuli will result in faster conditioning to A than to B, and the consequent reduction in α_B will interfere with further conditioning to B. Similar arguments would apply to the other examples of relative validity studied by Wagner (1969).

There is, however, one serious problem with this analysis. There is evidence that a less valid stimulus will not only fail to gain associative strength but also may apparently lose such strength as it had acquired during an earlier history of reinforcement. Wagner et al. (1968) showed that a light accompanied by a pair of tones would acquire significant control over responding, but that when the correlation of the tones with reinforcement was increased, the light lost control, even though its own correlation with reinforcement was unchanged. Similarly, if B is more salient than A, it may initially acquire control, even though it is a less valid predictor of reinforcement. As training continues, however, the difference in validity outweighs the difference in salience, and B loses control while A gains control (e.g., Jenkins, 1973; Rescorla, 1972b). An analysis that accounts for overshadowing by saying that a less valid stimulus fails to acquire associative strength because of a decrease in α is in no position to explain how it can lose strength that it has earlier acquired.

The simplest explanation of such results, consistent with the present theory, would be to assume that α affects both learning and performance. Even if its associative strength remained high, then, a stimulus might lose control over responding if a decline in α decreased the probability that this associative strength would be translated into performance. If B acquires control over responding, but is then presented only in conjunction with a second stimulus (A) better correlated with rein-

forcement, any decline in control by B may reflect the eventual decline in α_B brought about by the relative change in B's validity. It is too early to say whether this is a satisfactory solution; as noted later, this general case provides the only compelling reason for assuming that α affects performance as well as learning.

The present theory states that blocking depends on a decline in α to the added element; as such, blocking should never be complete and will depend on the rate of change in α . The most obvious factor determining the degree of blocking will be the extent to which the added stimulus signals some change in reinforcement.

If A alone signals reinforcement and AB signals nonreinforcement, then α_B will tend to increase rather than decrease, since B must be a better predictor of nonreinforcement than A. If AB signals a stronger reinforcer than that signaled by A alone, then α_B will tend to decrease, but less rapidly than it would if AB signaled the same reinforcer as A. This should presumably be sufficient to predict the unblocking observed by Kamin (1969) in his experiment in varying shock intensity, especially when it is remembered that in his experiment AB signaled a 4 mA shock, and this would have been sufficient to produce substantial levels of conditioning to B in one or two trials. There is, however, reason to believe that unblocking may be produced by the occurrence of any surprising event shortly after the presentation of the compound stimulus. Kamin (1969) reported that the delivery of a second shock 5 sec after each compound trial resulted in significant conditioning to the added element, even though this double shock did not appear to act as a stronger reinforcer. Gray and Apignanesi (1973) observed a similar effect when they presented a brief auditory and visual stimulus shortly after each compound trial. If confirmed, such results would suggest that unblocking was not simply due to an increase in λ enabling the added element to acquire associative strength, in the manner suggested by Rescorla and Wagner (1972). Unblocking

may occur whenever the added element predicts some event of consequence (such as the second shock in Kamin's experiment), whether or not that event is itself able to support conditioning. According to the present analysis, one could argue that the function of the surprising event was to prevent a decline in the value of α associated with the added element, thus enabling conditioning between that element and the original unconditioned stimulus (UCS) to proceed normally.

The application of the present rules to the case of acquired distinctiveness and dimensional transfer in instrumental discrimination learning is relatively simple. Relevant stimuli in a discrimination problem predict the occurrence and omission of reinforcement more accurately than any other stimuli, and an increase in their α values will follow from Equation 4. Similarly, irrelevant stimuli, uncorrelated with the delivery of reinforcement, are presented in conjunction with predictive relevant stimuli and should suffer a decline in α value.

The prediction of the effects of nonreinforced preexposure to a CS and of uncorrelated presentations of a CS and UCS requires that such a CS be regarded as forming a compound with a set of background stimuli, C. The crucial point, then, is that the presentation of AC (CS plus background) predicts no change from the probability of reinforcement or nonreinforcement predicted by C alone, and a decline in α_A should follow. There is, however, some question whether this is an ideal account. In the first place, it is only the phenomenon of latent inhibition that necessitates the otherwise rather unhappy assumption that α_A declines even when $|\lambda - V_A|$ is equal to $|\lambda - V_X|$. For reasons of theoretical symmetry, if for no others, one would expect this equality to produce no change in α_A . In the second place, there is reason to believe that random presentations of a CS and UCS retard subsequent conditioning more seriously than does simple nonreinforced preexposure to the CS; this additional effect, moreover, may be relatively specific

to conditioning with that particular UCS (Mackintosh, 1973). One interpretation of this observation is that α is both stimulus- and reinforcer-specific. Changes in the associability of a tone and shock, resulting from exposure to particular correlations between the tone and shock, will produce only generalized changes in the associability of that tone with other reinforcers.

If further research were to confirm this suggestion, the notion of α as a stimulus-specific learning-rate parameter would need replacing with a parameter $\alpha_{A,R}$, the value of which determined the magnitude of changes in the associative strength of A when it was paired with reinforcer R (or the omission of that reinforcer). If α thus represented the associability of a particular stimulus with a particular reinforcer, changes in α would be produced only by exposure to that stimulus and reinforcer (or similar ones), and the phenomenon of latent inhibition would lie outside the scope of the theory—to be explained, perhaps, in terms of some simpler mechanism of habituation.

In the absence of further evidence, further speculation along these lines would be idle. For the present, the rules suggested in Equations 4 and 5 may be regarded as one way of representing the idea that the associability of a stimulus with reinforcement will be affected by the predictiveness of that stimulus relative to that of other, concurrently presented stimuli. These rules are certainly not to be thought of as final, but rather as illustrating that it is possible to express this informal idea reasonably precisely. These particular rules, it will be recognized, bear a more than passing resemblance to Rescorla and Wagner's (1972) rules for changing V : In their system, a change in V_A depends on the status of both V_A and V_X ; here a change in α_A depends on the relative status of V_A and V_X . Informally, they assume that only if a reinforcer is otherwise unpredicted will its presentation in conjunction with A increase the associative strength of A; here it is assumed that only if a reinforcer is otherwise unpredicted will its presentation increase the probability of attending to A.

Starting Value of α : Representation of Stimulus Saliency

It is not enough to devise rules for changing α . A comprehensive theory must specify the limits within which α may change and the conditions determining the starting value of α . The simplest solutions to these problems are to allow α , as a learning-rate parameter, to vary between 0 and 1 and to assume that the starting value of α for any particular stimulus is positively correlated with the intensity or saliency of that stimulus. While these are reasonable assumptions, they are not in fact sufficient to account for the effects of stimulus saliency.

Lovejoy (1968), in his formalization of a two-stage theory of selective attention, distinguished between the fixed, base-level distinctiveness of a cue and its changeable or directable distinctiveness. Only the latter was subject to modification by learning. The implication was that changes in attention to various cues due to experience are "bounded, second-order effects which take place within definite limits that depend on the initial strength of the cues" (p. 60). The type of consideration that led Lovejoy to this solution was the observation that however much one might pretrain rats to attend to brightness cues in a jumping stand, they are unlikely to learn a simultaneous brightness discrimination as rapidly as a spatial discrimination. If the difference between spatial and brightness discriminations was represented only by differences in the starting value of a parameter such as α , and if sufficient exposure to a correlation between a set of stimuli and reinforcement resulted in the α value for those stimuli tending to 1.0, it would follow that differences in the difficulty of visual and spatial discriminations would eventually disappear. It seems clear that differences in the difficulty of instrumental discriminations cannot be adequately represented in this simple way; animals may never perform with complete accuracy if the discrimination is hard enough (e.g., Hara & Warren, 1961), and Mackintosh (1969), for example, noted that differences in the speed of reversing

an easy or difficult brightness discrimination after extensive overtraining were not predicted by such a model.

It is equally easy to show that such an account is inadequate for the data of classical conditioning. Although the main effect of differences in the intensity of the CS is a difference in the rate of conditioning, Kamin (1965) has shown that a very weak CS may result in a lower level of conditioning even at asymptote. Asymptotic differences in performance, of course, cannot be attributed to variations in a learning-rate parameter.

Differences in the discriminability of S+ and S- in a discrimination problem or between the presence and absence of the CS in classical conditioning, therefore, seem to have more permanent effects on performance than can be encompassed by a model that represents such differences by assigning different values to α , but which then permits α for any stimulus to vary between 0 and 1 as a result of experience. Instead of introducing a further parameter (such as a fixed α') for each set of stimuli, to represent their discriminability from the background or from other discriminative stimuli, it would seem more consistent with other psychological knowledge to rely on the notion of generalization. Perkins (1953) and Logan (1954), for example, argued that the well-authenticated observation that conditioning proceeds more rapidly with an intense CS than with one less intense may be a consequence of differences in generalization between the CS and the background. The more intense a CS, the greater the difference between that CS and its absence (the background). Reinforcement in the presence of an intense CS and nonreinforcement in its absence, therefore, will result in less generalization of excitation to the background and less generalization of inhibition from the background to the CS than will a comparable amount of differential reinforcement correlated with the presence and absence of a less intense CS. By the same line of argument, the greater the difference between S+ and S- in a discrimination problem, the greater the generalization of excitation to

S- that will result from each reinforcement of S+, and the greater the generalization of inhibition to S+ that will result from each nonreinforcement of S- (cf. Kendler, 1971).

These ideas can be formally represented in a relatively simple way by introducing a parameter, $S_{i,j}$, representing the similarity of the i th and j th stimuli, such that $0 \leq S \leq 1$. With two discriminative stimuli, A_1 and A_2 , reinforcement of A_1 will result in an increment in the associative strength of A_1 according to the usual equation:

$$\Delta V_{A_1} = \alpha_{A_1} \theta (\lambda - V_{A_1}). \quad (2)$$

By generalization from A_1 to A_2 , however, this trial will also result in an increment in the associative strength of A_2 , according to the equation,

$$\Delta V_{A_2} = S_{A_1, A_2} \alpha_{A_1} \theta (\lambda - V_{A_1}). \quad (6)$$

Similarly, nonreinforcement of A_2 will decrease the associative strength of A_2 , and will also result in a proportional decrease in the associative strength of A_1 .

A sufficiently high value of S_{A_1, A_2} will ensure that discriminative performance between A_1 and A_2 will never be perfect, just as a sufficiently high value of $S_{A, \bar{A}}$ (strong generalization between A and its absence) will ensure that even at asymptote the level of conditioning maintained by a weak CS will be below that maintained by a more intense CS. With suitable performance rules, intermediate levels of generalization will not affect performance at asymptote, but the generalization parameter will still provide an additional factor (over and above α) determining *rate* of learning. It is possible, therefore, to allow that with sufficient training the value of α associated with any stimulus can approach 1, without thereby abolishing differences in rate of learning correlated with differences in intensity or discriminability.

Specificity of α

Most formal theories of selective attention have followed Zeaman and House (1963) and Sutherland (1964) in assuming

that subjects learn to attend not just to specific stimuli, such as a vertical line, black door, or red response key, but to stimulus dimensions, such as line orientation, brightness, or hue. Sutherland and Mackintosh (1971), for example, assumed that changes in attention could be represented by changes in the strengths of various analyzers, each of which detected variations along a particular stimulus dimension. The assumption is a rather natural one for theories designed to account for the data of experiments on simultaneous discrimination learning, where subjects are usually confronted on each trial with a pair of stimuli, such as vertical and horizontal lines or black and white doors, which can clearly be described as differing along such dimensions. As Jenkins and Sainsbury (1969) have noted, however, it may not seem quite so naturally suggested by a successive discrimination between the presence and absence of a specific stimulus, and it can be applied only with some difficulty and little success to certain of the data obtained in such a situation.

There are, of course, experimental data that seem to provide good evidence for the dimensionality of attention. In studies of intradimensional and extradimensional shifts, the specific stimuli to which the subject is exposed change from one stage of the experiment to another; the observation that learning of the shift problem is faster if the relevant stimuli of the first problem differed along the same dimension as those of the shift certainly suggests that the changes in attention resulting from original training were changes in attention to entire stimulus dimensions. Similarly, if the phenomenon of transfer along a continuum (Lawrence, 1952) is to be interpreted in attentional terms, the implication is that an increase in the probability of attending to stimuli differing widely along a particular dimension results in an increase in the probability of attending to another set of stimuli differing less widely along the same dimension.

It is questionable, however, whether data such as these require the assumption that

analyzers appropriate to entire stimulus dimensions are strengthened and weakened. One could equally well assume that an increase in attention to one stimulus generalized to other stimuli in proportion to their similarity to the training stimulus. Thus an increase in α_{A_1} , resulting from the correlation of reinforcement with the presence of A_1 and nonreinforcement with its absence, might result in graded increases in α_{A_2} , α_{A_3} . . . α_{A_n} . Similarly, if A_1 signaled reinforcement and A_4 signaled nonreinforcement, as in an experiment on discrimination learning, generalized changes in α would accrue to the intervening stimuli A_2 and A_3 from both A_1 and A_4 .

This approach has the advantage of again narrowing the discrepancy between traditional, single-stage theories of associative learning and two-stage theories of selective attention. No appeal is made to hypothetical observing responses or analyzers detecting dimensions of stimulus change. A change in the state of one stimulus, whether of its associative strength, V , or of its learning rate parameter, α , is said to generalize to other stimuli in accordance with entirely traditional (even if unspecified) assumptions. There is the further advantage that this revision may explain the relatively small differences often observed in comparisons of intradimensional and extradimensional shifts in animal subjects. If changes in α to one pair of stimuli differing along a particular dimension result in only *generalized* changes in α to another pair, it is entirely reasonable to expect these latter changes to be rather small.

Performance Rules

Theories of learning are inferences from observed behavior; some relationship must be postulated between the theorist's terms and the experimental subject's responses. It will be sufficient, for present purposes, to suppose that the value of V_A bears some monotonic relationship to the probability that A , as a CS, will elicit an appropriate CR, or as a discriminative stimulus, will result in the occurrence of

an appropriate instrumental response. Although deliberately vague, this is neither contentious nor problematic. Theories of attention, however, are faced with the necessity of specifying further relations between their theoretical terms and their subjects' responses. As noted earlier, in traditional two-stage theories the direction of the subject's attention is assumed to affect not only what the subject learns as a consequence of the outcome of the trial, but also what response the subject makes on that trial. The probability that a rat will choose the black rather than the white arm of a T maze depends not only on the relative associative strengths of black and white, but also on the relative strengths of attention to brightness and other stimuli. In the original one-look model of Zeaman and House (1963), the subject attends to only one stimulus dimension on each trial, and behavior on that trial is controlled *only* by the response tendencies to those stimuli. Sutherland and Mackintosh (1971) adopted a more complex solution, assuming that behavior was sometimes controlled by only one analyzer, but that sometimes more than one set of stimuli would determine behavior.

There is, in fact, definitive evidence that in two-choice discriminations behavior can be controlled simultaneously by several sets of stimuli; rats can, after all, solve conditional discriminations. But there is equally good evidence pointing to this conclusion from studies of simple simultaneous discriminations, where solution could in principle be based on only one set of stimuli at a time. When more than one set of stimuli is simultaneously (but redundantly) relevant, it is possible to show that performance may be controlled by both on a single trial (Sutherland & Mackintosh, 1971, p. 142). Even in the case where one set of stimuli is relevant and another irrelevant, there is evidence that both relevant and irrelevant stimuli may combine to control performance. This is suggested by the fact that rats, trained on a simultaneous visual discrimination, having responded consistently to

one position over a long series of trials, may make no errors once they break their position habit (Sutherland & Mackintosh, 1971, p. 91). Analysis of the performance of animals shifted from one discrimination to a second, in which the originally relevant stimuli become irrelevant, has shown that few if any errors occur on those trials when the new S+ is combined with the former, but now irrelevant, S+. Such results suggest that both stimuli contribute to the subjects' choice behavior (Medin, 1973; Tighe, 1973).

These conclusions are quite consistent with the general trend of the present argument. It may be assumed, therefore, that all stimuli contribute to performance in accordance with their current associative strengths. The question remains whether one should assume that changes in α not only determine the rate of learning about specific stimuli but also affect the extent to which a stimulus' associative strength is actually translated into performance. One could certainly assume that the contribution of any stimulus to performance was a product of its associative strength and α value. Thus as α_A declined, even if V_A had some moderate value, the contribution of A to performance would also decline.

There is actually very little evidence bearing on this question. Since, in general, a stimulus with a low value of α will also be one with little associative strength, it may be very difficult to decide whether the reason why a particular stimulus fails to affect behavior is because it is not attended to (i.e., has a low α value) or because its associative strength is low. There is evidence from studies of discrimination learning that differences in the associative strength of S+ and S- from an initial discrimination are preserved while the animal learns another problem with the initial stimuli no longer correlated with reinforcement—even though there may be no detectable evidence that such stimuli affect behavior (e.g., latency of responding) on the intervening problem (Mackintosh, 1963; Stettner, 1965). Similarly, experiments on generalization in

pigeons have established that in the presence of one stimulus (e.g., a colored background on the response key), another set of stimuli, such as the orientation of a line on the key, may exert little or no control over responding, in spite of the fact that a particular line has been established as a signal for reinforcement, and reliably sloping gradients may be observed if different lines are presented on a black background (Newman & Benefeld, 1968).

Such instances of masking of control, however, can be explained without recourse to the assumption of competition between stimuli for control. The associative strength of the masking stimuli may be sufficiently great that ceiling effects obscure the control being exercised by the masked stimuli. These results, therefore, cannot really be said to provide conclusive evidence of the influence of α on control. Until such evidence is provided, it would seem reasonable to suggest that α may simply be a learning-rate parameter, with no effect on the control of behavior. Behavior would then be determined by a combination of the associative strengths of all stimuli present and impinging on the subject's receptors.

The only set of results that may require a revision of this assumption is the observation, noted earlier, that a stimulus imperfectly correlated with reinforcement may initially acquire associative strength but will eventually lose control if it is accompanied by other, more valid stimuli. The explanation provided by Rescorla and Wagner (1972) is that such a stimulus loses associative strength due to competition with more valid stimuli for a limited total amount of associative strength. The present theory is better able to explain why an invalid stimulus should not acquire associative strength in the first place than to explain why, having acquired strength, it should then lose it. It may be necessary to assume that although such a stimulus maintains its associative strength, a decline in α will decrease the probability that it will control responding.

CONCLUSIONS

The theoretical ideas proposed here may be summarized briefly. Changes in the associative strength of a stimulus are partly determined by a learning-rate parameter, α , specific to that stimulus. This parameter is itself subject to change, increasing when a stimulus predicts a change in reinforcement, decreasing when it does not. This idea is formally equivalent to one of the main tenets of two-stage, attentional theories of learning, namely, the assumption that the probability of attending to a stimulus determines the probability of learning about that stimulus and may itself change with experience. This equivalence may justify characterizing the present set of ideas as a theory of attention, but since that term has a number of connotations, it might be better to stress that what I am proposing is a theory about the associability of stimuli with reinforcement. At the risk of belaboring the obvious, I must reiterate that α is a learning-rate parameter, and possibly a determinant of performance: There is no implication that changes in α correspond to changes in perception.

Theories of selective attention have all incorporated a second assumption—that the probability of attending to one set of stimuli is inversely related to the probability of attending to others. The empirical justification for this assumption has been that it predicts the occurrence of such instances of stimulus selection as overshadowing and blocking. A closer analysis suggests, however, that these phenomena may not be due to any such competition for attention, but may rather be a consequence of subjects learning to ignore stimuli that signal only the occurrence of a reinforcer already predicted by other stimuli. This idea can be expressed by the choice of appropriate rules for changing α . Thus the single assumption that α increases and decreases in particular ways is sufficient to explain many of the data that have been thought to require a theory of selective attention.

It is worth concluding with some qualifications. It is obvious enough that the

ideas proposed here are more a program for a theory than a fully elaborated formal model of conditioning and discrimination learning. In our present state of knowledge, there is little need to apologize for this. It is more important to point to some factors that may limit the generality of the present arguments.

First, I have assumed that it is possible to predict the behavior of a subject toward a compound stimulus, AB, solely from knowledge of the status of its elements, A and B. The possibility of "configural" conditioning or "compounding" shows that this is not always true: With sufficient training, animals can learn to respond appropriately when AB is consistently reinforced but A and B separately are consistently not reinforced (e.g., Rescorla, 1972a; Woodbury, 1943). There is good reason to believe, moreover, that configural learning occurs even when it is not explicitly reinforced. Booth and Hammond (1971), for example, showed that continued reinforcement of AB might be sufficient to reduce significantly the level of responding maintained by A and B alone, even without nonreinforcement of the component stimuli in isolation.

Similar configural learning is a fairly common observation in operant discrimination learning by pigeons. If pigeons are rewarded for pecking an illuminated response key in the presence of a 1,000 Hz tone but not in its absence, they will learn not to respond to the key light alone but will equally refuse to peck in the presence of the tone if the key light is turned off (Jenkins & Harrison, 1960). This case is of particular interest in the present context, for it is an example of the main unresolved problem for the analysis proposed here—how a salient but relatively invalid stimulus comes to *lose* control over responding after it has initially acquired such control. If pigeons stop pecking when the key light is turned off, it is obviously misleading to suggest that the light has lost control over responding. Perhaps similar configural effects occur in other, similar situations. There are, in fact, very few data available to show how

far the apparent absence of control by B following reinforcement of AB reflects configural conditioning to the AB compound.

A second factor ignored by the present analysis may also contribute to at least some cases where salient irrelevant stimuli lose control of responding. From the time of their earliest discovery, results showing apparently selective effects in conditioning and discrimination learning have been attributed by some investigators to the occurrence of overt orienting responses. If B fails to acquire control over responding following reinforcement of AB, it is said to be because the subject oriented towards A and thus failed adequately to observe B. One of the few instances where there is actually good reason to suppose that such effects might operate is in the experiments of Jenkins and Sainsbury (1969), in which pigeons were trained to peck at visual displays containing discrete elements. In this situation, A and B would refer, for example, to small red and green dots or to a circle and a star; these elements were displayed on a large, specially constructed, response key, which enabled the experimenters to record which element was pecked at on a given trial. If birds were trained on a successive discrimination between AB+ and B-, Jenkins and Sainsbury found that they learned the discrimination by concentrating their pecks at A on reinforced trials. This suggested the possibility that when pecking at (orienting towards) A on reinforced trials, the pigeon associated only A with reinforcement. Although nominally reinforced on 50% of trials, therefore, B lost control of responding because it failed to gain strength on reinforced trials. Although later finding that pigeons could, to some extent, associate B with the outcome of a trial even though pecking at A, Jenkins (1973) also obtained results that suggest that the (partial) selectivity of orienting responses in this type of situation may indeed play an important role in enabling birds to learn the discrimination between AB+ and B-. If in such a problem reinforcement was forfeited on positive trials whenever the birds pecked at A,

they continued to peck at B on both positive and negative trials and thus failed to learn the discrimination. Because pecks were directed at B, it gained strength on reinforced trials and thus continued to control responding. It is possible, therefore, that in this situation at least, the loss of control by the relatively invalid stimulus, B, depends on a mechanism of selective orientation that lies outside the scope of the present analysis.

Finally, it is necessary to acknowledge the possibility that there may be situations in which selective effects occur for the reasons postulated by traditional theories of selective attention. Mackintosh (1971), for example, found that significant overshadowing of a weak tone by a strong light might occur on the very first trial of conditioning. Such overshadowing clearly cannot depend on subjects learning the redundancy of the tone nor on the mechanism proposed by Rescorla and Wagner (1972); it suggests a perceptual or attentional interaction. It is possible that similar instances of attentional interactions would be observed more frequently if they were sought under more appropriate conditions. I earlier argued that the inverse hypothesis of theories of selective attention did not rest on any very secure rationale: Animals may have a limited capacity for processing information, but it is hard to believe that this limitation prevents the simultaneous analysis of the relatively salient stimuli typically used in studies of conditioning. If there is a limit, however, its effects should be apparent under suitable conditions, such as the brief controlled presentations of complex stimuli.

This argument has also been advanced by Riley and Leith (in press), and its force is suggested by the results of an experiment, undertaken by Turner (Note 2), on delayed matching to sample in pigeons. Given unlimited exposure to the sample key, subjects performed accurately even when they did not know until after the sample had been turned off whether they were going to be required to match the color or the line displayed on the sample. Under these conditions, therefore,

they were clearly able to attend to both features of the sample. But as soon as exposure to the sample was reduced to a fixed, very brief period of time, performance deteriorated sharply, unless the subjects were given an additional, conditional cue that signaled whether the ensuing trial was going to require a line match or a color match. The implication is that pigeons were unable to analyze (or remember) both features of the compound sample when it was presented for only a brief interval. Provided the conditional cue signified which feature they needed to attend to, however, they maintained accurate performance. There was evidence, therefore, that stimuli were competing for access to a limited capacity system, and the inverse hypothesis was fully supported.

Although it may seem unpleasantly complex, it is possible that a complete analysis of conditioning and discrimination learning will require the assumption that the degree of competition between stimuli may vary from one extreme, where all available stimuli are analyzed on a single trial, to the other, where something like a one-look model may apply.

REFERENCE NOTES

1. Hall, G. *Learning to ignore irrelevant stimuli: Variations within and between displays*. Unpublished manuscript, 1974. (Available from G. Hall, Laboratory of Experimental Psychology, University of Sussex, Brighton, England.)
2. Turner, C. Personal communication, 1969.
3. Mackintosh, N. J. *Blocking of conditioned suppression: Role of the first compound trial*. Manuscript submitted for publication, 1975.

REFERENCES

- Booth, J. J., & Hammond, L. J. Configural conditioning: Greater fear in rats to compound than component through overtraining of the compound. *Journal of Experimental Psychology*, 1971, 87, 255-262.
- Egger, M. D., & Miller, N. E. When is a reward reinforcing?: An experimental study of the information hypothesis. *Journal of Comparative and Physiological Psychology*, 1963, 56, 132-137.
- Estes, W. K. Toward a statistical theory of learning. *Psychological Review*, 1950, 57, 94-107.
- Estes, W. K. The statistical approach to learning theory. In S. Koch (Ed.), *Psychology: A study of a science* (Vol. 2). New York: McGraw-Hill, 1959.

- Fisher, M. A., & Zeaman, D. An attention-retention theory of retardate discrimination learning. In N. R. Ellis (Ed.), *International review of research in mental retardation*. New York: Academic Press, 1973.
- Force, D. D., & LoLordo, V. M. Attention in the pigeon: Differential effects of food-getting versus shock-avoidance procedures. *Journal of Comparative and Physiological Psychology*, 1973, 85, 551-558.
- Gray, T., & Appignanesi, A. A. Compound conditioning: Elimination of the blocking effect. *Learning and Motivation*, 1973, 4, 374-380.
- Halgren, C. R. Latent inhibition in rats: Associative or nonassociative? *Journal of Comparative and Physiological Psychology*, 1974, 86, 74-78.
- Hara, K., & Warren, J. M. Stimulus additivity and dominance in discrimination performance by cats. *Journal of Comparative and Physiological Psychology*, 1961, 54, 86-90.
- Hull, C. L. *Principles of behavior*. New York: Appleton-Century-Crofts, 1943.
- Jenkins, H. M. Noticing and responding in a discrimination based on a distinguishing element. *Learning and Motivation*, 1973, 4, 115-137.
- Jenkins, H. M., & Harrison, R. H. Effect of discrimination training on auditory generalization. *Journal of Experimental Psychology*, 1960, 59, 246-253.
- Jenkins, H. M., & Sainsbury, R. S. The development of stimulus control through differential reinforcement. In N. J. Mackintosh & W. K. Honig (Eds.), *Fundamental issues in associative learning*. Halifax, Canada: Dalhousie University Press, 1969.
- Kamin, L. J. Temporal and intensity characteristics of the conditioned stimulus. In W. F. Prokasy (Ed.), *Classical conditioning: A symposium*. New York: Appleton-Century-Crofts, 1965.
- Kamin, L. J. Predictability, surprise, attention and conditioning. In B. Campbell & R. Church (Eds.), *Punishment and aversive behavior*. New York: Appleton-Century-Crofts, 1969.
- Kendler, T. S. Continuity theory and cue-dominance. In H. H. Kendler & J. T. Spence (Eds.), *Essays in neobehaviorism: A memorial volume to Kenneth W. Spence*. New York: Appleton-Century-Crofts, 1971.
- Krechevsky, I. "Hypotheses" in rats. *Psychological Review*, 1932, 39, 516-532.
- Lashley, K. S. *Brain mechanisms and intelligence: A quantitative study of injuries to the brain*. Chicago: University of Chicago Press, 1929.
- Lawrence, D. H. Acquired distinctiveness of cues: I. Transfer between discriminations on the basis of familiarity with the stimulus. *Journal of Experimental Psychology*, 1949, 39, 770-784.
- Lawrence, D. H. Acquired distinctiveness of cues: II. Selective association in a constant stimulus situation. *Journal of Experimental Psychology*, 1950, 40, 175-188.
- Lawrence, D. H. The transfer of a discrimination along a continuum. *Journal of Comparative and Physiological Psychology*, 1952, 45, 511-516.
- Logan, F. A. A note on stimulus intensity dynamism (V). *Psychological Review*, 1954, 61, 77-80.
- Lovejoy, E. *Attention in discrimination learning*. San Francisco: Holden-Day, 1968.
- Lovejoy, E., & Russell, D. G. Suppression of learning about a hard cue by the presence of an easy cue. *Psychonomic Science*, 1967, 8, 365-366.
- Lubow, R. E. Latent inhibition. *Psychological Bulletin*, 1973, 79, 398-407.
- Lubow, R. E., & Moore, A. U. Latent inhibition: The effect of nonreinforced preexposure to the conditioned stimulus. *Journal of Comparative and Physiological Psychology*, 1959, 52, 415-419.
- Mackintosh, N. J. Extinction of a discrimination habit as a function of overtraining. *Journal of Comparative and Physiological Psychology*, 1963, 56, 842-847.
- Mackintosh, N. J. Further analysis of the overtraining reversal effect. *Journal of Comparative and Physiological Psychology, Monograph*, 1969, 67 (2, Pt. 2).
- Mackintosh, N. J. An analysis of overshadowing and blocking. *Quarterly Journal of Experimental Psychology*, 1971, 23, 118-125.
- Mackintosh, N. J. Stimulus selection: Learning to ignore stimuli that predict no change in reinforcement. In R. A. Hinde & J. Stevenson-Hinde (Eds.), *Constraints on learning*. London: Academic Press, 1973.
- Mackintosh, N. J., & Holgate, V. Serial reversal training and nonreversal shift learning. *Journal of Comparative and Physiological Psychology*, 1969, 67, 89-93.
- Mackintosh, N. J., & Turner, C. Blocking as a function of novelty of CS and predictability of UCS. *Quarterly Journal of Experimental Psychology*, 1971, 23, 359-366.
- Medin, D. L. Subproblem analysis of discrimination shift learning. *Behavior Research Methods and Instrumentation*, 1973, 5, 332-336.
- Mellgren, R. L., & Ost, J. W. P. Transfer of Pavlovian differential conditioning to an operant discrimination. *Journal of Comparative and Physiological Psychology*, 1969, 67, 390-394.
- Miles, C. G., & Jenkins, H. M. Overshadowing in operant conditioning as a function of discriminability. *Learning and Motivation*, 1973, 4, 11-27.
- Newman, F. L., & Benefield, R. L. Stimulus control, cue utilization, and attention: Effects of discrimination training. *Journal of Comparative and Physiological Psychology*, 1968, 66, 101-104.
- Pavlov, I. P. *Conditioned reflexes*. London: Oxford University Press, 1927.
- Perkins, C. C., Jr. The relation between conditioned stimulus intensity and response strength. *Journal of Experimental Psychology*, 1953, 46, 225-231.
- Reiss, S., & Wagner, A. R. CS habituation produces a "latent inhibition effect" but no active "conditioned inhibition." *Learning and Motivation*, 1972, 3, 237-245.

- Rescorla, R. A. Summation and retardation tests of latent inhibition. *Journal of Comparative and Physiological Psychology*, 1971, 75, 77-81.
- Rescorla, R. A. "Configural" conditioning in discrete-trial bar pressing. *Journal of Comparative and Physiological Psychology*, 1972, 79, 307-317. (a)
- Rescorla, R. A. Informational variables in Pavlovian conditioning. In G. Bower (Ed.), *The psychology of learning and motivation* (Vol. 6). New York: Academic Press, 1972. (b)
- Rescorla, R. A., & Wagner, A. R. A theory of Pavlovian conditioning: Variations in the effectiveness of reinforcement and nonreinforcement. In A. H. Black & W. F. Prokasy (Eds.), *Classical conditioning II: Current research and theory*. New York: Appleton-Century-Crofts, 1972.
- Riley, D. A., & Leith, C. R. Multidimensional psychophysics and selective attention in animals. *Psychological Bulletin*, in press.
- Schnur, P. Selective attention: Effect of element preexposure on compound conditioning in rats. *Journal of Comparative and Physiological Psychology*, 1971, 76, 123-130.
- Shepp, B. E., & Eimas, P. D. Intradimensional and extradimensional shifts in the rat. *Journal of Comparative and Physiological Psychology*, 1964, 57, 357-361.
- Shepp, B. E., & Schrier, A. M. Consecutive intradimensional and extradimensional shifts in monkeys. *Journal of Comparative and Physiological Psychology*, 1969, 67, 199-203.
- Spence, K. W. The nature of discrimination learning in animals. *Psychological Review*, 1936, 43, 427-449.
- Spence, K. W. *Behavior theory and conditioning*. New Haven, Conn: Yale University Press, 1956.
- Stettner, L. J. Effect of prior reversal and elimination of inhibition on the persistence of a discrimination despite subsequent equal reinforcement of the discriminanda. *Journal of Comparative and Physiological Psychology*, 1965, 60, 262-264.
- Sutherland, N. S. The learning of discrimination by animals. *Endeavour*, 1964, 23, 69-78.
- Sutherland, N. S., & Andelman, L. Learning with one and two cues. *Psychonomic Science*, 1967, 15, 253-254.
- Sutherland, N. S., & Holgate, V. Two-cue discrimination learning in rats. *Journal of Comparative and Physiological Psychology*, 1966, 61, 198-207.
- Sutherland, N. S., & Mackintosh, N. J. *Mechanisms of animal discrimination learning*. New York: Academic Press, 1971.
- Thomas, D. R. Stimulus selection, attention, and related matters. In J. H. Reynierse (Ed.), *Current issues in animal learning*. Lincoln: University of Nebraska Press, 1970.
- Tighe, T. Subproblem analysis of discrimination learning. In G. Bower (Ed.), *The psychology of learning and motivation* (Vol. 7). New York: Academic Press, 1973.
- Trabasso, T. R., & Bower, G. H. *Attention in learning: Theory and research*. New York: Wiley, 1968.
- Turner, C. *Models of discrimination learning*. Unpublished doctoral dissertation, Oxford University, 1968.
- Wagner, A. R. Stimulus validity and stimulus selection in associative learning. In N. J. Mackintosh & W. K. Honig (Eds.), *Fundamental issues in associative learning*. Halifax, Canada: Dalhousie University Press, 1969.
- Wagner, A. R. Elementary associations. In H. H. Kendler & J. T. Spence (Eds.), *Essays in neobehaviorism: A memorial volume to Kenneth W. Spence*. New York: Appleton-Century-Crofts, 1971.
- Wagner, A. R., Logan, F. A., Haberlandt, K., & Price, T. Stimulus selection in animal discrimination learning. *Journal of Experimental Psychology*, 1968, 76, 171-180.
- Waller, T. G. Effect of irrelevant cues on discrimination acquisition and transfer in rats. *Journal of Comparative and Physiological Psychology*, 1970, 73, 477-480.
- Warren, J. M., Derdzinski, D., Hirayoshi, I., & Mumma, R. Some tests of attention theory with cats. In D. Mostofsky (Ed.), *Attention: Contemporary theory and analysis*. New York: Appleton-Century-Crofts, 1970.
- Woodbury, C. B. The learning of stimulus patterns by dogs. *Journal of Comparative Psychology*, 1943, 35, 29-40.
- Zeaman, D., & House, B. J. The role of attention in retardate discrimination learning. In N. R. Ellis (Ed.), *Handbook of mental deficiency: Psychological theory and research*. New York: McGraw-Hill, 1963.

(Received October 7, 1974)

An Emotion-Attribution Approach to Moral Behavior: Interfacing Cognitive and Avoidance Theories of Moral Development

Richard A. Dienstbier, Donald Hillman, John Lehnhoff,
Judith Hillman, and Maureen C. Valkenaar
University of Nebraska—Lincoln

A theory is presented concerning the impact of attributions about the causes of emotional responses as they influence self-control in temptation situations. Research is reviewed indicating a high level of adult sensitivity to external influence in making such causal attributions. Two studies are presented in which the posttransgression emotions of second-grade children are labeled *shame* (because of being found out) or *guilt* (due to the transgression itself); when a similar situation was subsequently represented as safe from detection, shame-condition children transgressed 60-80% more than guilt-condition subjects. It is suggested that emotional arousal elicited in temptation situations because of past punishment or options that are inconsistent with the self-image is *necessary* for inhibition but *not sufficient* unless attributed to a relevant cause. The literature on the relative effectiveness of moral socialization techniques is discussed with respect to the theory, and the relevance to cognitive dissonance and to overjustification approaches to motivation is discussed. An integration of social-learning and cognitive-developmental theories is approached through explicating the translation of moral decision into behavior by focusing on the ways that cognition may exert partial control over the impact of less finely differentiated emotional response, allowing cognitive overrides of contradictory emotional dispositions without eliminating the emotion.

Most of the modern literature on the development of conscience or morality is related to one of two major approaches—cognitive, concerned with moral reasoning and decision processes, or social learning, concerned with the emotional states and behaviors associated with self-control in the

face of temptation. Although both approaches are supported by the data of a great many studies, the two approaches remain essentially unintegrated. We seek to explicate some of the important steps in the translation of moral decision into moral behavior and to apply that analysis to an explanation of why certain moral socialization techniques are more conducive to internalization than others. Our theoretical perspective is presented in the early sections of the article; two supporting studies follow. The emphasis of the final section is on the relevance of the theory to traditional approaches to morality and to the socialization literature.

The theory can be stated in outline as follows: Behavior associated with self-control in the context of temptation is heavily influenced by negative emotional states usually characterized as anxiety, fear, guilt, and/or shame. This emotional mediation of moral behavior continues into adulthood, when the resultant behavior may relate to higher

The authors wish to thank Sister Leonarda and Richard Eisenhower for their kind permission to use the students and facilities of St. Teresa's School in Lincoln, Nebraska, and Norris School in Hickman, Nebraska, and to thank the administration of the Lincoln Public Schools, particularly Research Coordinator Terry Workman, for cooperation in securing subjects for the second study. Thanks are also due to Monte Page and Bruce Sales for their thoughtful review of the manuscript, to Shelley Stall and Jackie Kahl for their work in preliminary research as in-booth observers, and to Gail Zimmerman for her assistance in caring for the twins in Study 2.

Donald Hillman is now at the University of Vermont, and Judith Hillman is now at Johnson State College.

Requests for reprints should be sent to Richard A. Dienstbier, Psychology Department, University of Nebraska-Lincoln, Lincoln, Nebraska 68508.

orders of moral judgment as defined by modern cognitive approaches to morality. Although the association of these negative emotional states with decisional and behavioral processes may take place in part through mechanisms of conditioning and learning specified by traditional avoidance learning theorists (e.g., Mowrer, 1950), these associations depend heavily on the causal attributions that are made about the source of the negative emotions during socialization experiences. Social and situational influences (often temporary) play an important role in forming causal attributions; different physical and verbal socialization techniques particularly provide different information relevant to the child's causal attributions. Later, in the face of temptation, the impact of the emotion response on behavior will be heavily influenced by the still malleable beliefs held about the causal origins of the emotional response. When the individual believes detection of transgression is not possible, a negative emotional response in the face of temptation is necessary but not sufficient for the inhibition of transgression. In order for that emotional arousal to serve an inhibitory function, the individual must identify his emotional discomfort as due to a relevant cause, such as the transgression per se (and the implications of the transgression for self-image, etc.), rather than as due to an irrelevant cause, such as the fear of punishment. With maturity, control of emotional attributions passes from the socializing agent to the individual and becomes related to the level of moral development, although situational and social cues continue to play an important role.

EMOTION ATTRIBUTION

The work of Schachter and Singer (1962) demonstrated that when subjects injected with adrenalin did not anticipate that arousal effects would result, they tended to attribute their emotional state to the situation, feeling happier or angrier depending on the specific postinjection experimental context. That study was the first in a series by Schachter, his students, and others that indicated that people are quite flexible in the

manner in which they make causal inferences about the source and meaning of their emotional arousal.

Subsequent research exploring the implications of those findings for moral behavior followed the work of Lykken (1957). Lykken had shown that psychopathic individuals seemed not to experience normal emotion-induced inhibitions. Schachter and Latané (1964) demonstrated that psychopathic criminals (compared to normal criminals) would learn to avoid shocked errors (compared to nonshocked errors) in a lever-maze task only after sustaining arousal from an adrenalin injection. Schachter and Ono (cited in Schachter & Latané, 1964) found increased cheating by college students who had received the tranquilizer chlorpromazine (compared to placebo control groups). These studies can be viewed as evidence of the necessity of emotional arousal as an important component in the avoidance of responses with aversive consequences, with an increase in arousal facilitating that avoidance and a decrease attenuating avoidance. However, since only the degree of arousal was manipulated, and not interferences the subject might make concerning causal attribution, the necessity of arousal was demonstrated but the issue of the sufficiency of arousal was not addressed.

By illustrating the role of emotional arousal, these data provided support for Mowrer's (1950) two-factor avoidance model. Mowrer's model emphasized that the emotional response that the organism developed in response to the cues signaling an aversive event would motivate the avoidance response, which, when successful, would lead to the reinforcing reduction of the negative emotional state.¹ Emotional arousal in anticipation of punishment for a potential transgression should, by this interpretation, lead to the avoidance of the temptation behavior. But Schachter and Latané (1964) speculated that for the psychopath, the avoidance of behavior with aversive con-

¹ More modern theorists have extended this perspective, suggesting that anxiety becomes conditioned specifically to the "behavioral and cognitive precursors of the act" (Aronfreed, 1968, pp. 54-55).

sequences depended on more than the mere presence of arousal. Psychopaths experience a high level of physiological arousal with no apparent subjective emotional distress in response to cues that would evoke an emotional response in normal individuals. Schachter and Latané suggested that it was the chronic inability of the psychopath to experience his physiological arousal as emotion that accounted for his moral insensitivity.² These suggestions approach the perspective proposed in this article, but our view differs by suggesting that even in temptation situations where arousal is experienced as emotion, specific causal attributions are crucial in determining the behavioral outcome.

It was the Schachter and Latané (1964) research that led to a series of studies by the first author indicating that college-age men and women more often succumbed to the temptation to cheat when they could attribute the emotional arousal associated with cheating to a placebo pill, which supposedly had associated emotional side effects (Dienstbier, 1972; Dienstbier & Munter, 1971). Whereas Schachter and Singer (1962) had demonstrated that external cues are important in allowing the attribution of and experiencing of artificially induced arousal as specific emotion, the cheating research demonstrated the reverse—that naturally induced specific emotions could be interpreted as nonspecific arousal when subjects attributed the associated arousal to a nonemotionally relevant source such as the

placebo pill. Since these cheating studies are important for the theoretical perspective developed in this article, they are reviewed here in some detail.

Recruited as participants in a study on the effects of a vitamin supplement on vision, subjects were given a vocabulary test as a "delay task while waiting for the pill to take effect." Besides anticipating visual effects from the pill, a random half of the subjects were led to anticipate arousal effects of "a pounding heart, hand tremor, sweaty palms, a warm or flushed face, and a tight or sinking feeling in the stomach" (following Nisbett & Schachter, 1966). Control subjects anticipated benign side effects irrelevant to arousal. After the vocabulary test, to validate the potency of the pill, subjects viewed the auto-kinetic illusion, and were led to understand that the pill was responsible for their seeing the immobile light as moving. Thereafter, subjects were given the opportunity to check over their vocabulary test answers and to see the correct answers. They were informed that the department developing the vocabulary test was interested in using it for future prediction of college success, and that the "board of psychologists" responsible for it would like to interview any students who did particularly poorly on the test (less than 20 right). The test was designed so that less than 1% of college freshmen could attain that score without cheating.

The period during which the individual could look over the correct answers to the vocabulary test and his own answers was also introduced as a "delay period," in this case prior to the rating of side effects because "it was important that the subjects rate the side effects they might be experiencing just 9 minutes after they had experienced the visual effects." The salience of the side effects expectation was increased for that period by having subjects look over a side effects checklist at the onset of that period. Any cheating that took place during the delay period could be detected because pressure-sensitive paper (making a second copy of the answers) was placed under the answer paper during the taking of the test; the answer paper was removed by the subjects prior to the cheating period, so that there-

² Suggestions by Mandler (1964) and the theoretical structure proposed by Lazarus (1968) are useful in understanding possible mechanisms to account for the psychopath's lack of emotional experience in the face of arousal. Mandler suggested that the childhood of the typical psychopath is characterized by frequent emotional upheavals and that the individual might learn to minimize the emotional discomfort of these frequent traumas by ignoring the physiological signals from his body, which, if allowed to contribute to a strong emotional feeling, would greatly accentuate his distress.

Mandler's suggestions fit well with Lazarus's (1968) concept of benign reappraisal, the process by which the stress reaction is reduced by a cognitive reappraisal involving a lowered assessment of the potential danger in a situation after the individual has adjudged that he is powerless to affect the outcome.

after any changes made on the real answer paper would not correspond to those on the pressure-sensitive paper.

In the first study (Dienstbier & Munter, 1971) 49% of the subjects who anticipated arousal side effects from their placebo pills cheated, compared to 27% of the control condition, with cheating defined as changing any answers. This significant difference became more impressive when it was noted that the entire effect was due to the male subjects, whose cheating rates were 56% and 17% in the two conditions, respectively. Follow-up research (Dienstbier, 1972) indicated that women, too, respond to the arousal placebo with increased cheating if the threat of the consequences for failure is made less frightening (it appeared that in the original study the more intense anxiety experienced by women in response to the threat of being called by the "board of psychologists" interfered with their attending to the pill side effects). In a second study reported in the same article, it was further determined that the effect of the arousal placebo in reducing resistance to temptation was indeed due to the subject's attributing specifically to the pill—merely attending to arousal feelings without having taken the placebo did not produce the effect.

It appeared that all subjects experienced emotional arousal while confronting the temptation to cheat, largely as a result of considering the act of cheating and the potential negative consequences. But when subjects could attribute the cause of their arousal to the pill, that arousal ceased to have an inhibiting effect. It is not, in other words, that the subjects in either pill condition experienced more or less arousal, but rather that the attribution of the arousal to a neutral source (in the arousal pill condition) defused that arousal, preventing it from having emotional relevance.³

³ There are two lines of evidence suggesting that subjects make the attributions of emotional causality at a less than verbally conscious level. First, when subjects were interviewed, none ever indicated conscious awareness of such a decisional process. (Subjects were completely debriefed and questioned concerning possible suspiciousness following the cheating research.) Second, at the end

Whereas Schachter and Latané (1964) had speculated that this process in chronic form accounted for the moral insensitivity of the psychopath, the cheating studies suggested that the same process could be demonstrated to produce episodes of moral insensitivity in normal subjects.

The work of Solomon and Wynne (1954) and subsequent researchers emphasized that although emotional arousal was necessary in the acquisition of avoidance responses, the elimination of overt signs of arousal was often observed once such responses had been well learned. Except for the Schachter and Ono data (cited in Schachter & Latané, 1964), which demonstrated increased cheating with chlorpromazine, we would not have anticipated, therefore, that emotional arousal played any role in moral decision making for normal adults. Yet the cheating results suggested that even for 19-year-olds in temptation situations, the avoidance of transgression is still heavily dependent on emotional arousal. Additionally, the studies demonstrated that although 19-year-old adults should have had considerable experience in learning which features of their own behavior and of the environment lead to negative emotional arousal, they were surprisingly susceptible to external influences in attributing arousal experienced in the face of temptation.

But however powerful these data are, they do not provide the theoretical bridge to our final point. Our thesis requires that we demonstrate more than the defusing of emotional arousal through attribution to a non-emotional source in a context as artificial as that associated with the placebo pill studies. We must demonstrate instead that emotional

of the cheating period, when arousal subjects rated the degree to which they experienced the arousal side effects, the rate of reporting such side effects was surprisingly low. (There was, unfortunately, no control group with which they could be compared on this dimension, since the benign symptom subjects rated themselves on the benign symptoms.) It appears, therefore, that although the information concerning expected symptoms is given verbally, thereafter the very significant impact that these symptom expectations have upon cheating behavior is achieved by processes not apparent to the subject.

arousal may be defused for moral behavior even though still experienced as negative emotion, through the attribution to an emotion-inducing source that has no direct relevance to the moral decision.

SOCIALIZATION AND EMOTION ATTRIBUTION

We speculated that children would be even more receptive than adults had proven to be to cues that would suggest different causal attributions for emotional arousal, and that these cues would play an important role in determining the impact of socialization processes. Specifically, if a socialization pattern left the child with the impression that the cause of his emotional arousal or discomfort was his own misbehavior, the impact of this internal orientation on the child's behavior would be very different from the impact of an external orientation in which equally intense emotion was thought to be caused by fear of punishment. The latter attribution would be particularly ineffective in situations in which the tempted individual could be certain that detection was impossible.

Even early in their socialization history, most normal children will experience some emotional arousal following transgression and detection by a disapproving adult. Several features of the situation naturally lead to that arousal: the child's knowledge that the adult disapproves, previous scoldings now recalled, anxiety over the uncertainty of the outcome, and the likelihood that the adult will emphasize the discrepancy between the child's behavior and higher standards. The manner in which the socializing agent responds to the child's transgression will influence the causal attributions the child will make about his emotional discomfort. Socialization procedures that draw the child's attention to the transgression rather than to the aftermath (confrontation and punishment) facilitate the attribution of arousal to the act of transgression. Generally, such procedures involve discussions of the transgression in a manner that does not threaten the child, so that the child is able to dwell on the transgression itself and the behavioral standards he has violated, rather than on

the threat that confrontation offers. Often the self-concept is manipulated (e.g., "you are a bad boy") in relation to the standard or behavior ("because you didn't listen and lied"). The attribution of arousal to the anticipated transgression later leads the child to avoid the transgression itself. On the other hand, responses by socializing agents that threaten the child or that become salient through pain draw sufficient attention away from the transgression so that the child is likely to attribute his emotional discomfort exclusively to confrontation with the socializing agent. When later facing a similar temptation situation, a child treated in the latter manner might experience a high level of emotional arousal, but by attributing his arousal to fear of detection and punishment he would tend to resist temptation only if he believed that detection was likely.

In our research with children we wished to buttress our theory by establishing a temptation and transgression situation in which the emotional arousal following detection would be equal for the two groups of children, with the children assured that future transgressions could not be detected. We then imparted to the children an internal (due to own behavior) or an external (due to confrontation) causal attribution concerning the source of that emotional arousal. It was hypothesized that the children who believed that they "felt bad" for their previous transgression (internal orientation, or guilt) would misbehave far less than children in the same situation who believed that they "felt bad" because their previous transgression had been detected (external orientation, or shame).

STUDY 1

Method

In Study 1 of the two studies with children, second-grade subjects were brought, one at a time, to a "new toy room" in their school and were asked to keep watching a slot car go slowly around an oval track during the experimenter's absence; they were to stop the car with a control device if the car began to go too fast. Although the child was given sufficient justification for watching the car, other electrical toys "with loose wires" con-

trolled by a hidden observer⁴ were started and stopped according to a schedule to distract the child. If the child transgressed by allowing himself to be distracted from the implied duty of watching the slot car, the observer caused the slot car to jump the track (the event that was supposed to be avoided) and simultaneously signaled the experimenter to return.

Pretending that the slot car was broken, the experimenter asked the child to accompany him to a separate room, where the independent manipulation was administered: labeling the child's feelings as due to transgression (guilt) or due to being found out following the transgression (shame). The child was then returned to the room, taught to operate the slot car correctly, and assured that this "new" car would not break if it jumped the track; he was also assured that the door would now be locked and the experimenter did not have a key, implying that detection of transgression of failing to watch the slot car would be impossible. The dependent measure was the total time spent not attending to the slot car during this second watching phase of 12 minutes (with the experimenter absent). A reliability of .99 was obtained between two observers for five subjects; thereafter one observer was used.

One male and one female experimenter participated,⁵ each using boys and girls as subjects, so that with the independent manipulation of guilt versus shame, a $2 \times 2 \times 2$ factorial design was effected.

Results

Results indicated no significant effects for either the sex of experimenter or children, but the nonsignificant difference between the mean seconds transgressing for the shame condition (164) and the guilt condition (102) were suggestive. A further analysis indicated that most of the demonstrated increased effectiveness for the guilt manipulation was due to the female experimenter—187 seconds transgressing in the shame condition, compared to 80 seconds in the guilt condition, $t(22) = 1.91$, $p < .07$, two tailed. It was our impression that a real difference did exist, despite the lack of a significant main effect for the shame versus guilt conditions or an experimenter by condition interaction. It was suspected that the differences between experimenters were due to their different experience levels with children (the more effective one was an elementary school

teacher and a parent) rather than a true sex difference.

The inconclusiveness of the findings of Study 1 led to the changes incorporated into the second study, which will be presented in somewhat greater detail.

STUDY 2

In an effort to reduce the variations in background between children in the two conditions and to allow a matched-pair type of analysis, second-grade twins were recruited by seeking the cooperation of their parents and of the twins themselves. Except for some minor changes in procedure and script and the participation of only one experimenter (the first author, who fit the criterion of having sufficient experience with children), the design of the study was essentially the same as for Study 1. Children were not considered separately by sex, since the results of Study 1 suggested no sex differences.

Method

The first twin to participate was chosen randomly and brought to the toy room while the other twin played quietly with an assistant in the waiting room. The participating twin was seated in front of the booth and was told that the slot car presently on the track was very old so it might break if it did fall off the track, but that the problem with the cars sometimes going too fast was in the track. The experimenter's first absence was so that he could "go to call the man who fixes the toys." The scheduled use of distraction toys such as the electric train "with the loose wires" was essentially the same as in the first phase of Study 1. Although the criterion time before the slot car would jump the track and the booth light would go on for the first twin was 6 continuous seconds of distraction, if the child had not met that criterion by the end of the 6th minute, the criterion time was reduced by 1 second for each passing minute of attentiveness of the child. Since the members of each twin pair tended to act remarkably alike, their first phase participation times tended to be very similar.

Immediately following the child's transgression, the experimenter returned. Upon "discovering" the transgression and the "broken" slot car, the experimenter asked the child to accompany him to the adjoining room, where he identified the twin's experimental condition and administered the appropriate manipulation according to the following script (the sections in parentheses indicate the additions that change the basic guilt script to the shame script, or where two sets of parentheses oc-

⁴ A role filled by the second and fifth authors.

⁵ A role filled by the third and fourth authors.

cur together, the first is the shame section and the second the guilt section:

I bet you feel a little bad now that (I found out) the car fell off. I've seen other kids feel bad when (someone found out) they weren't able to do exactly what they were supposed to do. When the other kids who tried to watch the car couldn't, they felt bad (when I found out) too.

But it's important that we do stop the car when the light first comes on, and before it flies off the track, so we'll try again with this new car. You won't feel bad if (I find this car still) (this car stays) on the track—if (I find) you've done a good job, you'll feel good, won't you?

Have you noticed that when other kids (can show people that they) have done the right thing, that they feel good? And when (people find out) they did things they were not supposed to, they feel bad, don't they? (When someone finds out) (Even if no one ever finds out) they feel bad for not doing what they were supposed to do, don't they?

After returning to the toy room, the child was directly informed that since he now was being taught to start and control the new unbreakable car, and since the door would be locked and the experimenter had no key, that "even if it did fly off the track I'd never find out, since you could put it back by yourself and I can't come right in with the door locked." Most children spontaneously acknowledged that this was true. With that statement, the experimenter signaled the observer⁶ to put his hand over his ears while the final part of the independent manipulation was administered: "Remember, you'd feel bad if (I ever found out you did a bad job) (you did a bad job, even if I never found out). There is no way that I could find out this time, though." The experimenter then left the room, locking the door and very loudly checking it to make sure it was adequately locked.

A 12-minute period, spiced with the same distracting toys starting up and the same dependent measure (the time spent not attending to the slot cars) was identical to the second phase of Study 1.

Results and Conclusions

Of the nine pairs of twins who came to the laboratory, complete data were gathered for six pairs. Two pairs were lost because one of the twins failed to follow instructions by shutting off the slot car so that he would not have to watch it. One pair was not run because the twins were hyperactive and inattentive. For each of the six pairs run, the twin receiving the shame instructions transgressed almost twice as much as the twin in

⁶ Thanks are due Steve Slane for his faithful and competent filling of the observer role.

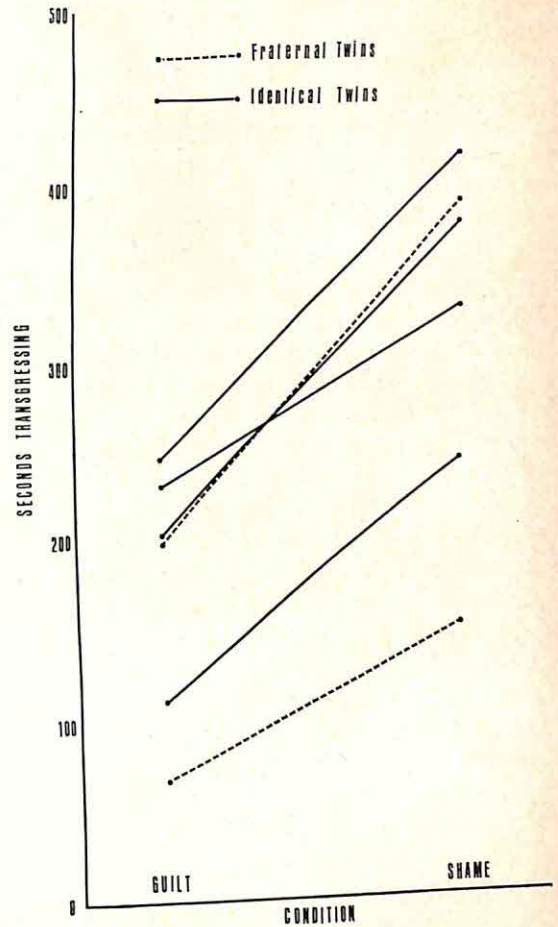


FIGURE 1. Seconds transgressing for the members of each pair of twins according to condition.

the guilt condition. The average time transgressing was 322 seconds for the shame condition, and 177 seconds for the guilt condition $t(5) = 8.0$, $p < .001$, two tailed. The regularity of the effect across pairs of twins is illustrated by Figure 1.

It is apparent that 8-year-old children are sufficiently cognitively sophisticated so that the differences between the guilt and shame manipulations are perceived, even in the face of the anxiety elicited by the confrontation with a mildly disapproving and strange adult. Additionally, the findings of 60% (Study 1) and 80% (Study 2) increases in transgression time for shame subjects over guilt subjects support conclusions from the adult cheating studies—the attribution of emotional causality has a large impact on the

behavior of children in resistance-to-temptation situations.

The degree to which the two child studies are acceptable as powerful support of the theoretical development is dependent on several assumptions that cannot be empirically demonstrated. Although we endeavored to maintain equal levels of emotional arousal for the children in the two conditions, and although there are no apparent reasons why the shame or guilt script variations might have induced different arousal levels, no direct assessment of emotional arousal was made. Also, we cannot be certain that the greater effectiveness of the guilt script was not due to one of the scripts drawing the subject's attention to considerations irrelevant to the attribution of emotional arousal. It might have appeared, for example, that the experimenter was less interested in the behavior of the children in the shame condition. Although the similarity of the scripts was meant to preclude this possibility, it is impossible to be certain that the script differences did not trigger different reactions between the children in the two conditions. However, since neither of these two explanations could easily be used to explain the results of the cheating studies with adults, the hypothesis involving the causal attributions of emotional arousal seems the only likely explanation common to both research approaches. Although the following discussion assumes that the child studies provide strong empirical support for the theoretical developments, we believe that it is possible to follow and accept the theoretical aspects of our thesis without being totally convinced of the relevance of the data from the child studies.

One possible approach to explaining the high dependence of our child and adult subjects on external cues rather than past experience in making causal attributions of emotional arousal is to suggest that this flexibility is greatest when the emotional state in question is an uncomfortable one. Some support-by-default for this view comes from the emotion-attribution literature; the impact of causal attribution manipulations has been convincingly demonstrated only with the negative emotions of anxiety, fear, and

startle response (Dienstbier, 1972; Dienstbier & Munter, 1971; Nisbett & Schachter, 1966; Ross, Rodin, & Zimbardo, 1969; Storms & Nisbett, 1970; Zanna & Cooper, 1974). In the first author's research, several failures to observe the causal reattribution of arousal associated with emotions not related to fear reinforced this view. Further support is provided by the research of Lazarus and his associates (Lazarus, 1968), who have demonstrated a readiness in their subjects to use cognitive information that allows the reduction of stress.

Another approach to understanding the apparent ease with which the children accepted the attribution of arousal to their own behavior or internal causes rather than to detection or external causes is to focus on the readiness of children in this age category to accept such internal attributions. Wolman, Lewis, and King (1971) demonstrated that as children pass from "young" (ages 5-7) through "middle" (ages 8-9) to "old" (ages 10-13) categories, they report that the conditions that stimulate emotions more frequently come from within themselves. Kagan (1971) has summarized a wide range of literature suggesting that the child's capacity for guilt increases most rapidly between the ages of 4 and 10.

RELEVANCE TO MAJOR THEORETICAL APPROACHES

Social Learning Approaches

Influenced heavily by traditional learning theory on the one hand, and by the psychoanalytic concept of identification on the other, the social learning approach to morality has progressed in the direction of the analyses of those features of the model that facilitate identification and those features of discipline techniques that facilitate resistance to temptation and appropriate affect.

Since the impact of modeling and identification research appears to be waning, we will not attempt an integration of our view with that branch of the social learning literature.⁷

⁷ The modeling literature, typified by the work of Bandura and his associates (Bandura, 1968) and Sears and his associates (Sears, Maccoby, & Levin, 1957; Sears, Rau, & Alpert, 1965), has never freed itself from the issues of identification versus simple

Instead, our discussion will be concerned with socialization techniques and the impact of those techniques on the affective and cognitive responses that facilitate resistance to temptation. Several issues in the social learning and the socialization-of-morality literature will be viewed from our theoretical perspective.

Punishment. As a behavior modification technique, and as demonstrated in many laboratory studies, punishment (including physical) is often a viable control technique. That punishment is effective in behavioral control (so that its various parameters bear analysis) is seldom questioned, and those laboratory studies that have compared punished and nonpunished control groups have generally demonstrated the superiority of punishment in behavioral suppression (see Johnston, 1972, for a recent review). But the socialization literature based on naturalistic observation in home and school settings indicates that the use of physical punishment is negatively associated with the internalization of social control (Bacon, Child, & Barry, 1963; MacKinnon, 1938; Maurer, 1974; Nowicki & Segal, 1974; Sears, Maccoby, & Levin, 1957; Sears, Rau, & Albert, 1965).

The dilemma of physical or power-oriented punishment being successful in the laboratory and counterproductive in less controlled settings, particularly as compared to induction, relates to our approach according to the following analysis. When all other things are equal, the more severe the punishment, the more the emotional arousal and the more the resistance to deviation in the future, *unless* the more intense punishment causes different attributions about the source of the emotional discomfort associated with that situation or distracts from other administered messages relevant to emotional

attribution. The first part of that statement is an affirmation of classical avoidance theory—that greater arousal will lead to more resistance to deviation. But, as is also suggested, intense physical punishment should indeed cause different attributions of arousal from either mild punishment or other disciplinary techniques that do not direct attention away from the transgression. If the physically punished individual faces future temptation in situations in which detection seems unlikely, then we would predict a paradoxical effect, with more intense punishment or more physically oriented punishment leading to less resistance to temptation.

Punishment that is love oriented, however, seems less destructive in its impact on the development of resistance to temptation. There are several reasons why such techniques are less counterproductive than physical discipline and why they are found to be positive in some studies. While physical punishment is usually accompanied by parental expressions of anger, encouraging similar expressions in the child, love-withdrawal techniques are associated with anger inhibition (Hoffman, 1970) in the child; the experiencing of anger by the child suggests that emotional arousal is being attributed to confrontation with the parent. Additionally, the love-withdrawal techniques tend to leave the child in a state of emotional discomfort for longer periods, with the termination of such punishment often dependent on some action or statement by the child that relates to the child's initial transgression. Expressions such as "I'm sorry that I did that" are emotional attribution statements frequently encouraged by parents in psychological punishment situations prior to the termination of the punishment. That love-oriented techniques have sometimes been found to be most effective when practiced by warm parents (i.e., Sears et al., 1957) may be due to the child's finding it more difficult to attribute his discomfort to the warm parents' behavior, rather than to his own. The fact that with greater warmth, there is more love to withdraw in the punishment situation is also relevant.

imitation, and in the 1970s seems to have lost its appeal as a research area. The failure of many major studies in this area to find convincing evidence of causal links between identification and resistance to temptations (i.e., Sears et al., 1965) has led the author of one significant recent review to suggest that the evidence linking morality and identification was meager at best (Hoffman, 1970). While the disinhibiting effects of a transgressing model seemed established, the inhibiting effects of a self-denying model were not.

Due to the unfamiliarity and associated uncertainty that the young subject experiences in the laboratory, combined with (usually, in our view) insufficient assurances of the undetectability of deviance, the laboratory provides neither a setting in which the exclusive attribution of arousal to punishment is likely nor one in which feelings of total freedom from detection are likely. In laboratory studies the punished child therefore usually appears more self-controlled than the nonpunished control. Unfortunately, a complete resolution of this point is not possible by reference to the morality literature, since researchers seldom attempt to provide evidence that their children were convinced that their transgressions could not be detected. Given that the laboratory is often quite unfamiliar to the research subjects, for young children to be secure in their belief of undetectability considerable effort should be devoted to the staging of experimental situations; if possible, the impossibility of detection should be articulated to the child (as in Study 2).

Naturalistic settings often permit the child to make causal attributions that negate the effect of punishment and lead to freedom from detection. In a home or classroom the harshly or frequently punished child has repeated opportunities to make accurate attributions that the negative emotions he experiences are due to confrontation and punishment rather than transgression and to become certain about when detection is possible. Those accurate emotional attributions facilitated by repeated experience in natural settings are not conducive to internalization. With the development of accurate attributions about the source of negative emotional arousal, either of two types of results may obtain (probably most such situations result in some of each of these two processes). The first process has been emphasized by those social learning theorists who have been heavily influenced by the traditional avoidance model exemplified by Mowrer's (1950) work, but who must explain the data that indicate the ineffectiveness of punishment in natural settings. Simply stated, that view suggests that the punished child learns accurate discriminatory cues so that no arousal

results during transgressions that occur in the absence of punishing agents. While it is apparent that such a process may take place, we would argue that there are a great many temptation situations for normal individuals in which arousal is elicited even though the individual is fully aware that detection is impossible; both the cheating studies and our child studies provided situations designed to make detection appear impossible, yet arousal and the causal attributions made about that arousal played a key role in determining the outcome.

We seek what we perceive to be a necessary middle ground between older theories, which emphasize automatic conditioning procedures to the exclusion of cognitive processes, and more modern approaches, which emphasize total cognitive control over the learning of contingencies. The older view (Mowrer, 1950), suggesting a classical-conditioning-like mechanism to account for the attachment of negative emotions to situations and behavior, could not account for data such as that offered in the cheating studies or the child studies, in which cognitive interpretations given to subjects relevant to the meaning of their emotional arousal had important effects. More modern approaches, such as that of Bindra (1974), suggest that reinforcing or punishing events do not directly affect the strength of the preceding response; rather what is learned are the "contingent relations" existing between the events of the situation. Bandura (1969, p. 553) suggests a similar relationship in stating that the classical-conditioning-like aversive techniques used in behavior modification procedures affect the client only to the extent that he is *consciously* willing to recall and revisualize or reexperience the therapy situation when faced with temptation in later outside situations.

Our view suggests a greater degree of autonomy between cognition and emotion, with cognition influenced by, but not under the total control of emotional arousal, and emotional arousal likewise influenced by, but not completely controlled by cognition interpretations. Thus the cognitive knowledge possessed by particularly the very young child that his emotional upset in the face of

temptation stems directly from previous punishment may not eliminate that emotion when detection is thought to be impossible; rather, such a causal attribution will make the emotion irrelevant and will allow the child to act somewhat as if the emotion were not present. To illustrate this cognition-emotion interaction, consider behavior in the face of a mild phobia. The fear experienced by an acrophobic adult while climbing an open stairway may be as intense when the fear is cognitively perceived to be unjustified (the stairs are safe) as when it is perceived to be justified (the stairs may break); yet through moderate fear levels, such cognitions will determine whether the individual will override his emotional arousal and continue to climb. We present the acrophobia illustration as more than a metaphor and suggest that many of the more important moral prohibitions learned by individuals with well-developed consciences have a phobialike character, with emotional arousal elicited even though the tempted individual may cognitively appraise the situation as perfectly safe from detection. Such an appraisal allows the emotion to be overridden without being eliminated if the individual attributes his arousal to a fear of detection or to other causes that may be irrelevant if detection is impossible.⁸

Recent research by Loftis and Ross (1974) suggests the possibility of a second effect resulting from different causal attributions. After the acquisition of an emotional response, the extinction rate for that response may depend in part on the source to which the individual retrospectively attributes his emotional arousal. In their research, a tone was heard while a shock-induced emotional reaction, as measured by galvanic skin response (GSR), was conditioned to a light. When later instructions

indicated that the emotional arousal could be attributed to the tone, extinction of the GSR was faster (compared to a control group for which the tone was not described as arousing), even though no tone was used during extinction. Thus, although our view emphasizes that emotional arousal will have differential effects depending on the causal attributions made, we do not exclude the process of different attributions eventually affecting arousal level. With respect to moral behavior, the two processes should tend to influence responses similarly, with attributions that arousal is due to irrelevant sources (such as fear of punishment when detection is not possible) both negating the effect of the arousal and, with repeated experience, leading to a reduction in emotional arousal as well.

Techniques such as induction (Hoffman, 1970), which do not depend on punishment, will undoubtedly result in less negative emotional arousal associated with transgression than techniques based on punishment; but successful induction, involving the explanation to the child of why certain behaviors are right or wrong, focuses attention on the behavior per se, so that the arousal that is present during such socialization episodes will be attributed to the act of transgression rather than to the threat of punishment. Hoffman's review indicates that induction is more effective in moral socialization than punishment-based techniques, even including love-oriented punishment. Our emphasis on the necessity of arousal for induction to be effective is relevant to Aronfreed's (1968, p. 9) suggestion that "available evidence usually points to great discrepancies between children's verbal expression of evaluative standards and their actual behavior in a real social context." The data of Henshel (1971), indicating increasing correlations (from fourth to seventh grade) between honesty assessed verbally and resistance to temptation in a cheating task, provide the basis for an explanation of the discontinuity between verbal expression and moral behavior noted by Aronfreed. Henshel suggested that behavior is not dependent as much on simply "knowing" what is right as it is on "feeling what is 'desirable.'" Unless

⁸ An extreme example of an apparently similar process is found in the work of Fenz and Epstein (Epstein, 1967) with skydivers. Although emotional arousal as indicated by physiological monitoring remains high as the skydiver gains experience, his view of the sport as not posing objective danger allows the enthusiast to jump. The changing attributions fostered by experience allow the experience of intense excitement where previously fear had dominated.

standards become associated with emotional responses, children will be as free to violate these standards as were the subjects in the cheating studies who could attribute their arousal to a pill.

Guilt and resistance to temptation. A second dilemma from the social learning literature is that although a correspondence is assumed (in most versions of social learning theory) between guilt (usually as measured by projective techniques) and resistance to temptation, some studies have failed to show such a correspondence (Allinsmith, 1960; Burton, Maccoby, & Allinsmith, 1961; Maccoby, 1959).

Some authors (e.g., Hoffman, 1970) have suggested that one of the major problems with that literature is that researchers have seldom endeavored to measure guilt in relation to the same types of transgression on which the behavioral measures are taken. Furthermore, the difficulty of interpreting projective measures of guilt is often cited. Our approach lends support to that explanation and allows a further articulation of the importance of measuring guilt and resistance to temptation with respect to the same transgression. As demonstrated by the cheating and child research, our approach emphasizes the high degree of dependency on immediate situational cues in attributing causation of emotional arousal. Thus it is quite likely that individuals who make causal attributions about the internal origins of arousal (guilt) in one situation will find that either their own past experience or current situational cues lead them to make more external attributions concerning emotional causality (shame or fear of punishment) in other temptation situations. Significant correlations between guilt and resistance to temptation should, therefore, be expected only when those dimensions are compared with respect to the same behavior. The only study that measured guilt and resistance to temptation with respect to the same type of transgression found that a significant relationship did exist (MacKinnon, 1938).

Cognitive Dissonance and Self-Perception Theory

The changing of the value of certain behaviors and objects through unnecessarily

strong negative external inducement (threat) has been demonstrated by researchers investigating a hypothesis derived from cognitive dissonance theory. That hypothesis stated that children who receive a mild threat to inhibit an act will take more personal responsibility for resisting temptation than will severely threatened children; they will therefore experience more cognitive dissonance over resisting temptation. Aronson and Carlsmith (1963) and Freedman (1965) demonstrated that when children were given a mild threat to not play with a valued toy, they valued the forbidden toy less (Aronson & Carlsmith) and played with it less (Freedman) than was the case under severe threat. The sophisticated control conditions of the Freedman study (finding no threat-effect differences when there was no opportunity to play with the toy during the threat-relevant period) allowed the elimination of many potential explanations alternative to the dissonance-derived hypothesis.

A more recent literature has developed indicating similar effects for unnecessarily strong positive inducements (reinforcement). The thinking of deCharms (1968) on the importance of perceived locus of control in motivation led to research on the overjustification effect by Deci (1971, 1972), Kruglanski, Alon, and Lewis (1972), Lepper, Green, and Nisbett (1973). This research indicated that when an individual is offered reinforcement for a task he would have performed without payment, he acts (in future situations without payment) as if he were less intrinsically motivated than before the unnecessary payment.

Both the dissonance and the overjustification research provide support for Bem's (1972) self-perception theory—that individuals tend to consider all available evidence in assessing the motivation underlying their own behavior, just as an observer looks for all the forces relevant to another's behavior in making attributions about the other's motivation. Our approach suggests that the overjustification and dissonance results are due to the effect that an offer of tangible reinforcement or threat of severe and tangible punishment has on reducing attributions

about the internal cause of positive and negative emotions. But by focusing as it does on the ability to change causal attributions of emotion associated with and resulting from behavior after the behavior has been performed, our approach suggests that reinforcement or punishment given after the fact will often create effects similar to inducements and threats made before the fact. This conclusion from our studies with children is reinforced by the findings on "retrospective misattribution" by Loftis and Ross (1974) cited earlier. The importance of this extension becomes more obvious when we consider that socialization efforts are usually made after some motivational failure has occurred.

As suggested earlier, the extent to which punishment or reinforcement draws attention to itself rather than to the preceding behavior will be influenced not only by the specific form of punishment or reinforcement but also by the style of administration and the situation. Thus while tangible reinforcers, such as physical punishment, are likely to draw the attribution of emotional response toward themselves rather than to the preceding behavior, certain approaches may accentuate that process. Information that suggests that the administered reinforcement should be sufficient compensation and an adequate inducement for future performance should increase the likelihood that pleasure derived from the task itself (from the means, as well as from the completion) will be attributed instead to the reinforcement.

This approach may be useful in illuminating an apparent contradiction in the recent overjustification literature. Unexpected prizes administered to a group of game-winning students by Kruglanski et al. (1972) had the effect of reducing liking for those competitive games (compared to nonrewarded winners), while children who received unexpected prizes for good drawing in the Lepper et al. (1973) study did not later behave as if their prizes caused them to devalue the special drawing task used. From our perspective, the critical difference lies in the changed attributions resulting from Kruglanski et al. telling their subjects

that they had been promised the rewards prior to the games (which in fact was not true) and in giving tangible reinforcers (plastic puzzles), which were salient during the game ratings because they were given before the rating. The positive emotional response resulting from playing and winning the games was apparently partially attributed to the prizes. In the Lepper et al. study, however, the unexpected reward group was not misled to believe that their rewards had been previously promised nor were those rewards salient at the time the dependent measure was taken; their awards had been posted (several days previously) on an "award board" that was not present in the classroom in which the dependent measure was taken.

Our approach leads to a theoretical integration between findings associated with achievement motivation and the overjustification research, an integration very similar to that which is apparent between the punishment literature (in natural settings) and the dissonance literature. Research in the achievement area indicates that while gestures of parental love and affection (nontangible reinforcers that do not draw attention from the child's behavior) are related to high achievement motivation in boys, the administration of tangible reinforcers by parents is not associated with high achievement motivation (Rosen & D'Andrade, 1959; Spence, 1970; Winterbottom, cited in McClelland, Atkinson, Clark & Lowell, 1953).

THE COGNITIVE AND AVOIDANCE THEORY INTERFACE

Cognitive development theories of morality, as exemplified by the writing of Piaget and Kohlberg, are too extensive to be reviewed here in detail. Rather, cognitive developmental theory will be briefly outlined and the relevance of that approach to the proposed theory will be explicated at a general level.

Piaget (1932) postulated that the development of the child's moral judgments corresponds to the development of other more general cognitive capacities. Prior to the development of sufficient cognitive flexibility, the child makes moral judgments in strictest

conformity to established rules, viewing behavior as totally right or wrong, basing that assessment on the outcome of an act regardless of the intent of the actor, and believing in the concept of "immanent justice"—that punishment will automatically and surely follow transgression. With maturity and cognitive flexibility, moral judgments become more flexible, with rules seen as changeable depending on other constraints of the situation; the outcome of a specific act has less impact on the assignment of blame than the intent of the actor. In part, this transition to flexibility is made possible through active participation in society and through rule making and changing, allowing rules to be seen as serving man's needs. Social participation also encourages the acquisition of increased empathy through interaction with peers.

Kohlberg (1958) extended Piaget's approach to a 6-stage sequence, each stage of which is defined by a different basis for moral judgment. In earliest stages, avoidance of punishment (Stage 1) and later the definition of right and wrong based on the valence of the outcome for the actor (Stage 2) prevail. Middle stages involve judgments based on the maintenance of social approval (Stage 3) and then on duty to laws and the prevailing social order (Stage 4). Higher stages are based on an understanding of the social-contract nature of social and legal systems and respect for the rights of others (Stage 5) and finally a recognition that legal and moral principles are often in conflict, resolved through principles of conscience, logical consistency, and mutual trust in fulfilling higher values (Stage 6).

Neither Piaget nor Kohlberg has extensively addressed theoretical questions of the relationship of moral behavior to moral reasoning, or of the relationship of affect and emotional arousal to moral reasoning. Kohlberg (1969), however, approached our viewpoint by suggesting that affective responses would progress developmentally with cognitive responses and that for the higher moral stages, guilt would be associated with different types of transgression. Thus for a Stage 6 individual, "guilt over violation of internal principles" would result; for the

Stage 4 individual, guilt would imply "concern about one's responsibility according to rules"; for the Stage 3 individual, only shame would be experienced.

Kohlberg's astute analysis is completely compatible with our orientation, but it is through the emotion-attribution perspective that we can address the issue of the dynamics that account for this presumed usual compatibility of affect and moral judgment in the maturing individual. The key to our explanation is that temptation situations are basically unchanging as the individual matures through the cognitive stages postulated by the cognitive-developmental theorists. That is, even though perceived somewhat differently with maturity, and even though different responses become available, the basic issues that lead to the emotional arousal associated with temptation remain essentially the same. It is, of course, impossible to prove this point, but it is difficult to name any moral issue faced by a mature adult in any sphere of personal or professional activity that does not have an obvious and basic counterpart for a 6-year-old child. There are mature and immature forms of stealing, of being dishonest or disloyal, of hurting others or being unfair, or of being destructive to property and life. This view is supported by the moral dilemmas used in the Kohlberg (1958) system to assess maturity of moral judgment—the situations can be seen as posing ethical dilemmas to individuals at any stage of moral maturity.

How does this continuity of moral dilemmas provide a mechanism for understanding the (normally) continuing compatibility of moral judgment and its supportive affect as the individual's level of moral maturity changes? A basic tenet of our approach is that many moral dilemmas evoke a phobia-like response, causing emotional discomfort as the individual begins to seriously consider transgression as an option. But that emotional arousal will have different meanings (and will affect behavior differently) for individuals at different levels of moral development. Thus one's stage of moral development does not determine whether an individual will experience arousal, but rather once such arousal is experienced, one's stage

of moral maturity will influence what the arousal means and therefore whether the arousal is relevant to the available behavioral options. We do not wish to exclude the possibility that the level of maturity will influence one's perception of the situation and that this perception will influence the amount of arousal experienced, but the impact of such initial perceptions on arousal level may be minimal.

As an example, we might consider the problem of killing others in a war, assuming a society in which parental instruction, the major institutions of society, and the child's personal encounters with death all create in the child a negatively toned emotional arousal at the prospect of causing another's death. Other factors that might support this arousal include the recognition of the extraordinary exercise of power in killing, the irreversibility of death, and the degree to which the behavior is relevant to the support of or the negation of self-concept. When ordered to kill in war time, individuals at different stages in Kohlberg's (1958) moral development scheme would act differently. The Stage 4 individual, believing firmly in the supremacy of the law and the absolute need to uphold the law, would agree to kill, yet he would experience inhibiting negative emotional arousal (based on processes such as those discussed earlier), which cannot be eliminated readily through cognitive control. Unless overwhelmed with negative emotion, his attribution of affect to a fear of killing would be sufficient to make that negative affect less relevant, allowing him to kill (just as the shame-oriented child is able to transgress even when aroused if detection seems impossible, or as cheating is facilitated for the individual who can attribute his arousal to a placebo). With killing experience, the Stage 4 soldier's emotional response should decline relatively rapidly (since, as indicated by Loftis and Ross, 1974, the attribution of arousal to an irrelevant source may facilitate subsequent extinction).

Individuals at Kohlberg's Stage 6 would initially experience arousal for the same reasons and in a manner similar to the Stage 4 individual. But principles based on individual conscience allow more latitude of out-

come. One may hierarchically arrange personal values so that killing is required by conscience to defend principles more important than the deaths involved, or one may decide that killing cannot be justified by the principles defended by the war. In the first instance, the arousal would be experienced as guilt (the individual assumes personal responsibility for the decision) over the anticipation of killing, but the attribution of that arousal to the violation of moral principles would not prevent the killing—the arousal would be of little relevance since other moral principles associated with the nominal purpose of the war have become more important.

The Stage 6 individual who has concluded that life is more important than those principles defended by the military action, on the other hand, experiencing his arousal as anticipatory guilt over the killing, would focus on his arousal as relevant to that behavior. Therefore, the arousal could be experienced as a manifestation of the degree of opposition the individual feels toward killing and as evidence of the correctness of the decision not to kill. As the three hypothetical individuals approach their battlefield decisions, increasing arousal will therefore have different impacts on each. The Stage 4 individual, feeling his arousal as unwelcome fear, might begin to feel cowardly and ashamed of his emotional reaction. The first Stage 6 individual, feeling anticipatory guilt, would feel more confused, wondering if this increased arousal should be taken as evidence of the wrongness of killing and hence of a need for the reordering of those values that favored killing. The second Stage 6 individual would, of course, become more convinced of the correctness of his decision not to kill.

Emotional arousal and cognition are thus seen to interact in a complex manner, with behavioral outcome determined by the outcome of that interaction. When either is overwhelming, it may tend to dominate despite the counterinfluence of the other, but in the normal middle ground of life, that mechanism which permits us to display our humanity (when humanity is defined in the old-fashioned manner as the dominance of the rational over the appetitive and emo-

tional) seems to depend heavily on our ability to experience our emotion as inoperative by attributing it to causes not relevant to the present situation. Tomkins (1963) wrote that it is "one of the major functions of any negative affect theory to guide action so that negative affect is not experienced," and that this represented "affect at a distance" (pp. 320, 321). We would agree, but add that the major function of any theory that seeks to integrate cognition, affect, and behavior must provide an explanation of how cognition, more differentiated and complex than affective responses, may moderate the impact of emotional arousal on behavior, even while failing to completely dominate or eliminate these affective responses.

REFERENCES

- Allinsmith, W. The learning of moral standards. In D. R. Miller & G. E. Swanson (Eds.), *Inner conflict and defense*. New York: Holt, Rinehart & Winston, 1960.
- Aronfreed, J. *Conduct and conscience: The socialization of internalized control over behavior*. New York: Academic Press, 1968.
- Aronson, E., & Carlsmith, J. M., Effect of the severity of threat on the devaluation of forbidden behavior. *Journal of Abnormal and Social Psychology*, 1963, 66, 584-588.
- Bacon, M. K., Child, I. L., & Barry, H., III. A cross-cultural study of correlates of crime. *Journal of Abnormal and Social Psychology*, 1963, 66, 291-300.
- Bandura, A. Social-learning theory of identifiatory processes. In D. A. Goslin & D. C. Glass (Eds.), *Handbook of socialization theory and research*. Chicago: Rand McNally, 1968.
- Bandura, A. *Principles of behavior modification*. New York: Holt, Rinehart & Winston, 1969.
- Bem, D. J. Self-perception theory. In L. Berkowitz (Ed.), *Advances in experimental social psychology* (Vol. 6). New York: Academic Press, 1972.
- Bindra, D. A motivational view of learning, performance, and behavior modification. *Psychological Review*, 1974, 81, 199-213.
- Burton, R. V., Maccoby, R. R., & Allinsmith, W. Antecedents of resistance to temptation in four-year-old children. *Child Development*, 1961, 32, 689-710.
- deCharms, R. *Personal causation*. New York: Academic Press, 1968.
- Deci, E. L. Effects of externally mediated rewards on intrinsic motivation. *Journal of Personality and Social Psychology*, 1971, 18, 105-115.
- Deci, E. L. Intrinsic motivation, extrinsic reinforcement, and inequity. *Journal of Personality and Social Psychology*, 1972, 22, 113-120.
- Dienstbier, R. A. The role of anxiety and arousal attribution in cheating. *Journal of Experimental Social Psychology*, 1972, 8, 168-179.
- Dienstbier, R. A., & Munter, P. O. Cheating as a function of the labeling of natural arousal. *Journal of Personality and Social Psychology*, 1971, 17, 208-213.
- Epstein, S. M. Toward a unified theory of anxiety. In B. A. Maher (Ed.), *Progress in experimental personality research* (Vol. 4). New York: Academic Press, 1967.
- Freedman, J. L. Long-term behavioral effects of cognitive dissonance. *Journal of Experimental and Social Psychology*, 1965, 1, 145-155.
- Henshel, A. M. The relationship between values and behavior: A developmental hypothesis. *Child Development*, 1971, 42, 1997-2007.
- Hoffman, M. L. Moral development. In P. H. Mussen (Ed.), *Carmichael's manual of child development* (Vol. 2, 3rd ed.). New York: Wiley, 1970.
- Johnston, J. M. Punishment of human behavior. *American Psychologist*, 1972, 27, 1033-1054.
- Kagan, J. *Personality development*. New York: Harcourt Brace Jovanovich, 1971.
- Kohlberg, L. *The development of modes of moral thinking and choice in the years ten to sixteen*. Unpublished doctoral dissertation, University of Chicago, 1958.
- Kohlberg, L. Stage and sequence: The cognitive-developmental approach to socialization. In D. Goslin (Ed.), *Handbook of socialization theory and research*. Chicago: Rand McNally, 1969.
- Kruglanski, A. W., Alon, S., & Lewis, T. Retrospective misattribution and task enjoyment. *Journal of Experimental Social Psychology*, 1972, 8, 439-501.
- Lazarus, R. S. Emotions and adaptation: Conceptual and empirical relations. In W. J. Arnold (Ed.), *Nebraska Symposium on Motivation* (Vol. 16). Lincoln: University of Nebraska Press, 1968.
- Lepper, M. R., Green, D., & Nisbett, R. E. Undermining children's intrinsic interest with extrinsic reward: A test of the "overjustification" hypothesis. *Journal of Personality and Social Psychology*, 1973, 28, 129-138.
- Loftis, J., & Ross, L. Retrospective misattribution of a conditioned emotional response. *Journal of Personality and Social Psychology*, 1974, 30, 683-687.
- Lykken, D. T. A study of anxiety in the sociopathic personality. *Journal of Abnormal and Social Psychology*, 1957, 55, 6-10.
- Maccoby, E. E. The generality of moral behavior. *American Psychologist*, 1959, 14, 358.
- MacKinnon, D. W. Violation of prohibitions. In H. A. Murray, *Explorations in Personality*. New York: Oxford University Press, 1938.
- Mandler, G. Comments on Dr. Schachter's and Dr. Latané's paper. In D. Levine (Ed.), *Nebraska Symposium on Motivation* (Vol. 12). Lincoln: University of Nebraska Press, 1964.

- Maurer, A. Corporal punishment. *American Psychologist*, 1974, 29, 614-626.
- McClelland, D. C., Atkinson, J. W., Clark, R. A., & Lowell, E. L. *The achievement motive*. New York: Appleton-Century-Crofts, 1953.
- Mowrer, O. H. *Learning theory and personality dynamics*. New York: Ronald Press, 1950.
- Nisbett, R. E., & Schachter, S. Cognitive manipulation of pain. *Journal of Experimental Social Psychology*, 1966, 2, 227-236.
- Nowicki, S., Jr., & Segal, W. Perceived parental characteristics, locus of control orientation, and behavioral correlates of locus of control. *Developmental Psychology*, 1974, 10, 33-37.
- Piaget, J. *The moral judgment of the child* (M. Worden, trans.). New York: Harcourt, Brace, 1932.
- Rosen, B. C., & D'Andrade, R. G. The psychosocial origin of achievement motivation. *Sociometry*, 1959, 22, 185-218.
- Ross, L., Rodin, J., & Zimbardo, P. G. Toward an attribution therapy: The reduction of fear through induced cognitive emotional misattribution. *Journal of Personality and Social Psychology*, 1969, 12, 279-288.
- Schachter, S., & Latané, B. Crime, cognition, and the autonomic nervous system. In D. Levine (Ed.), *Nebraska Symposium on Motivation* (Vol. 12). Lincoln: University of Nebraska Press, 1964.
- Schachter, S., & Singer, J. E. Cognitive, social, and physiological determinants of emotional state. *Psychological Review*, 1962, 69, 379-399.
- Sears, R. R., Maccoby, E. E., & Levin, H. *Patterns of child rearing*. Evanston, Ill.: Row, Peterson, 1957.
- Sears, R. R., Rau, L., & Alpert, R. *Identification and child training*. Stanford, Calif.: Stanford University Press, 1965.
- Solomon, R. L., & Wynne, L. C. Traumatic avoidance learning: The principles of anxiety conservation and partial irreversibility. *Psychological Review*, 1954, 61, 353-385.
- Spence, J. T. The distracting effects of material reinforcers in the discrimination learning of lower and middle-class children. *Child Development*, 1970, 41, 103-112.
- Storms, M. D., & Nisbett, R. I. Insomnia and the attribution process. *Journal of Personality and Social Psychology*, 1970, 16, 329-337.
- Tomkins, S. S. *Affect, imagery, consciousness: Vol. II. The negative affects*. New York: Springer, 1963.
- Wolman, R. N., Lewis, W. C., & King, M. The development of the language of emotions: Conditions of emotional arousal. *Child Development*, 1971, 42, 1288-1293.
- Zanna, M. P., & Cooper, J. Dissonance and the pill: An attribution approach to studying the arousal properties of dissonance. *Journal of Personality and Social Psychology*, 1974, 29, 703-709.

(Received December 4, 1974)

THEORETICAL NOTE

A Testable Property of Some Discrete-State Models for Recognition Memory

Dirk Vorberg
*Universität Konstanz,
Konstanz, West Germany*

Rainer Schmidt
*Technische Universität Darmstadt,
Darmstadt, West Germany*

Several current theories of recognition memory combine the notion of a few discrete states of memory with the assumption of continuous distributions of familiarity or apparent oldness. If a test for threshold theories of signal detection proposed by Krantz is applied, it can be shown that the discrete-state memory models imply a lower bound on the a posteriori curve obtained from confidence ratings in yes-no recognition experiments. From this the prediction is derived that new items cannot be perfectly recognized as new unless the forgetting probability is zero. Experimental data are presented that show this prediction to be wrong.

Some years ago, Bernbach (1967) and Kintsch (1967) proposed models for recognition memory that combined the assumption of only a few discrete states of memory with that of continuous distributions of "apparent oldness" or "familiarity," on which a decision process operates. These models enjoy some popularity, particularly since some of the more comprehensive theories of memory, such as those of Bernbach (1970), Kintsch (1970b), and Anderson and Bower (1972), reduce to them under suitable conditions. On the other hand, it has been repeatedly claimed (e.g., Kintsch, 1970a; Murdock, 1974) that high-threshold models of recognition memory are rejected by the known experimental findings. These claims seemingly were reinforced by Krantz's (1969) article, which presents an important empirical test for threshold theories of signal detection. What seems to have been overlooked in recent research on recognition memory is that both the Bernbach and the Kintsch models can be regarded as generalized high-threshold models and consequently can be subjected to the same critical test. The pur-

pose of this article is to show how Krantz's test can be applied to examine the discrete-state assumption of the models and to review some critical experimental data.

This paper was written while the first author was a National Science Foundation postdoctoral fellow at the Department of Psychology, New York University. We are indebted to Doris Aaronson, Patrick Berg, William K. Estes, Jean-Claude Falmagne, Robert Hogan, Micha Razel, Joan Gay Snodgrass, and James Townsend for helpful comments on earlier versions of this paper.

Requests for reprints should be sent to Dirk Vorberg, Fachbereich Psychologie, Universität Konstanz, 775 Konstanz, Postfach 7733, West Germany.

A yes-no recognition experiment can be imagined in which old and new items are presented singly and the subject's task is to evaluate each item as old or new. In addition to the "yes, old" or "no, new" decision, the subject is required to give a rating indicating the degree of confidence in his response. We will concentrate on the so-called a posteriori curve, which plots the probability that a test item was in fact old for each rating category. We assume the rating scale to be ordered, running from highest confidence that the item was old to highest confidence that the item was new. If the subject's use of the rating scale is reasonable, a monotonically increasing a posteriori curve is expected; the probability that the item was old should start from small values for high-confidence new responses and rise to values near 1 for high-confidence old responses. This is what is usually found and is easily accounted for by any reasonable theory of recognition memory. If the discrete-state assumption holds, the a posteriori curve must have a lower bound, below which no values can fall except for sampling errors. That is, the probability that an item was old, given that the subject responded "new," cannot generally attain values of 0, even for high-confidence responses. Stated differently, according to these models subjects cannot recognize new items perfectly unless they have perfect memory.

In the following discussion, we derive the lower bound on the a posteriori curve predicted by the Bernbach (1967) and the Kintsch (1967) models and give an equivalent lower bound on the slope of the receiver operating characteristic (ROC) curve. First, we give a brief statement of the models. In order to avoid redundancy, we then derive the lower bound using a more general model. In the final section, the models' predictions are compared to data.

PROPOSED DISCRETE-STATES MODELS

Bernbach's (1967) model assumes that an item can be in one of two states: a memory state (R) and a forgotten or unlearned state (N). All new items are initially in state N . Old items at the time of testing are either in state R , with probability $1 - \delta$, or in state N , with probability δ , where δ denotes the forgetting probability. Attached to each state is a distribution of apparent oldness, with densities $f_R(x)$ and $f_N(x)$, respectively. As a consequence of the assumptions, the distribution for new items is given by $f_N(x)$, whereas that for old items is a mixture of the two distributions $f_R(x)$ and $f_N(x)$. The model for recognition memory proposed by Anderson and Bower (1972), although derived from different assumptions, is formally equivalent to Bernbach's model.

Kintsch's (1967) model is more complicated. It has three memory states and three distributions of familiarity, but there is no one-to-one correspondence between them. Rather, the memory states are defined by intervals on the familiarity scale. Since we are mainly concerned with the decision aspects and not the learning process as described by this model, we shall simply ignore the memory states in this article.

According to Kintsch's model, all new items have familiarities determined by a distribution $\psi(x)$. After presentation of an item, it is assigned a new familiarity value. If its initial familiarity was less than a critical value ξ , its new familiarity is determined by a distribution $\phi(x)$; if it was larger than ξ , the new distribution will be $\phi'(x)$, where $\phi'(x)$ is assumed to equal $\phi(x)$ truncated at ξ . An item may be forgotten with probability f if its familiarity was less than a second critical value $\lambda > \xi$, where forgetting means reassignment of a familiarity value according to the initial distribution $\psi(x)$. Summarizing the assumptions, we have a similar picture to that in Bernbach's model: New items follow a distribution $\psi(x)$, whereas the distribution of old items is a

mixture of the three distributions $\psi(x)$, $\phi(x)$ and $\phi'(x)$.

Let us now introduce a more general notation, which encompasses both models. Any given item can be in one of three states, D_j , $j = 1, 2, 3$. All new items are in state D_1 ; old items are either in state D_1 , D_2 , or D_3 . For brevity, we let d_j stand for the probability that an old item is in state D_j , $P(D_j|\text{old})$. Since any old item is in exactly one state, $\sum d_j = 1$; for Bernbach's model $d_3 = 0$, as there exist only two states in this model.

Let the confidence ratings be numbered from R_1 to R_n , where low numbers indicate high-confidence old responses. In order to derive the a posteriori curve, it is necessary to specify how the confidence rating given to an item is determined by the item's state. Both Bernbach's and Kintsch's models handle this by attaching continuous distributions to the states and by assuming that the subject adopts a set of cutoffs that divide the underlying scale into intervals; the subject responds with R_k if and only if the oldness or familiarity of the item presented falls into interval I_k . Since the property to be derived holds for any underlying distribution and any response rule, we do not follow this approach. Instead, we simply define a set of conditional probabilities $P(R_k|D_j)$, which gives the probability of responding with R_k if the item is in state D_j for all k and j . The only assumption necessary is that the probabilities of the various ratings depend only on the item's state. These conditional probabilities can be thought of as derived from the assumption of underlying distributions and multiple cutoffs, or they can simply be regarded as free response bias parameters, which are under the subject's control. For brevity, we let r_{kj} stand for $P(R_k|D_j)$; since exactly one response is given, we have

$$\sum_k r_{kj} = 1, \quad \text{for } j = 1, 2, 3.$$

If we let O and N stand for the events that a test item is old or new, respectively, the a posteriori curve is then defined as $P(O|R_k)$, $k = 1, n$; that is, the probability that the item is old given the subject's response is R_k . From Bayes' theorem we have

$$P(O|R_k) = \frac{P(R_k|O)P(O)}{P(R_k|O)P(O) + P(R_k|N)P(N)} \quad (1)$$

where $P(O)$ and $P(N)$ denote the experimentally controlled probability that a test item

is old or new, respectively. According to the assumptions

$$P(R_k|O) = \sum_{j=1}^3 P(R_k|D_j)P(D_j|O) \\ = r_{k1}d_1 + r_{k2}d_2 + r_{k3}d_3,$$

since any old item is in one of the three states, and the probability of using a given rating depends only upon the item's state. For new items we have

$$P(R_k|N) = P(R_k|D_1) = r_{k1},$$

since all new items are in state D_1 . If we insert this into Equation 1, the result is

$$P(O|R_k) = \frac{(r_{k1}d_1 + r_{k2}d_2 + r_{k3}d_3)P(O)}{(r_{k1}d_1 + r_{k2}d_2 + r_{k3}d_3)P(O) + r_{k1}P(N)}.$$

If we replace $P(N)$ with $1 - P(O)$ and divide¹ every term by r_{k1} , we find

$$P(O|R_k) = \frac{d_1P(O) + \left(\frac{r_{k2}}{r_{k1}}d_2 + \frac{r_{k3}}{r_{k1}}d_3\right)P(O)}{d_1P(O) + 1 - P(O) + \left(\frac{r_{k2}}{r_{k1}}d_2 + \frac{r_{k3}}{r_{k1}}d_3\right)P(O)} \quad (2)$$

It is straightforward to prove that the right side of Equation 2 is at least as large as $d_1P(O)/[d_1P(O) + 1 - P(O)]$, since the term

$$\left(\frac{r_{k2}}{r_{k1}}d_2 + \frac{r_{k3}}{r_{k1}}d_3\right)P(O)$$

is nonnegative. Thus, the final result is

$$P(O|R_k) \geq \frac{d_1P(O)}{d_1P(O) + 1 - P(O)} > 0 \quad \text{if } d_1, P(O) \neq 0. \quad (3)$$

Equation 3 shows that the lower bound on the a posteriori curve is determined by the probability that an old item is in state D_1 , d_1 , and the probability that a given test item is old, $P(O)$. Since $P(O)$ is nonzero in any reasonable recognition memory experiment, it follows that, according to the discrete-state models, the a posteriori curve can attain zero only if $d_1 = 0$, that is, old items are never in state D_1 . It should be noted that no assumption about the rating probabilities r_{kj} was needed to derive Equation 3. Therefore, the bound given in Equation 3 holds for any

¹ We are assuming $r_{k1} \neq 0$. If $r_{k1} = 0$, $P(O|R_k) = 1$, which is consistent with the bound given in Equation 3.

choice of the r_{kj} , or equivalently, any possible response assumption, in particular for those of Bernbach (1967) and Kintsch (1967).

The parameter d_1 is essentially the forgetting probability in the models, since all new items are initially in state D_1 and old items can be in state D_1 only if they returned from D_2 or D_3 , that is, are forgotten. For Bernbach's model, we immediately see that $d_1 = P(D_1|O) = \delta$, the forgetting parameter. Thus, this model implies that perfect recognition of new items as new, that is, $P(O|R_k) = 0$ for high-confidence new ratings R_k , occurs only if there is no forgetting.

For Kintsch's model, d_1 is derived as follows: g denotes the probability that the familiarity of a new item exceeds ξ , and $a \times c$ and c denote the probabilities that an item in state D_2 or D_3 , respectively, exceeds λ . Since forgetting, that is, return to state D_1 , occurs with probability f if the item's familiarity was less than λ , it follows that $d_1 = (1 - g) \times (1 - a \times c)f + g(1 - c)f$. This expression equals zero if $f = 0$ or if $g = 1$, $c = 1$, or $c = 1$, $a = 1$, which means that either there is no forgetting, or new items are transferred to a perfect memory state on the first presentation. The model thus makes the same prediction as Bernbach's: Perfect recognition of new items as new occurs only if no forgetting is possible.

There are two ways of testing the lower bound predicted by the discrete-state models. One is to check whether empirical a posteriori curves reach the zero level under conditions where the assumption of no forgetting can be rejected on other grounds. This will be demonstrated in the next section. It should be noted that failure to reject the discrete-state models does not lead to their acceptance, since continuous models such as strength theory (e.g., Norman & Wickelgren, 1969) can account for bounded a posteriori curves as well. Also, it might be that the obtained curves do not extend to zero simply because an asymptote was not reached in the experiment, possibly due to too few rating categories or insufficient motivation of the subjects to use the highest rating categories properly.

The second possibility of testing the discrete-states assumption is to use independent estimates of the forgetting probability, d_1 , derive a quantitative prediction for the asymptote of the a posteriori curve, and compare the prediction to data. For example, we know from short-term memory experiments employing the Peterson and Peterson (1959)

paradigm that a single consonant trigram (CCC) is recalled after 18 sec with a probability of about .2 (e.g., Murdock, 1961). If the recall rate could be used to estimate the recognition forgetting-probability, we predict a lower bound of $.8 \times .5 / (.8 \times .5 + .5) = .444$ for a corresponding recognition experiment, with $P(O) = .5$. Thus, under these conditions we should find that on at least 44% of those trials where the subject responded new with highest confidence, the test item was in fact old, which would be hardly better than chance.

The problem with this second approach is that acceptable independent estimates for the recognition forgetting-probability are difficult to find. According to the original version of Bernbach's (1967) model, all items in state R will be correctly recalled, so that equating recall and recognition forgetting-rates would be defensible except when guessing occurs during recall. However, in a more recent version of the model, Bernbach and Kupchak (1972) proposed that the recall probability equals the probability of an old item being in state R times a multiplier less than one, which can be interpreted as a retrieval probability. Similarly, the dual-process theory of recall and recognition (Kintsch 1970b; Müller, 1913) maintains that recall involves item retrieval and an implicit recognition decision, whereas recognition involves only the latter. Thus, according to these models, the recall forgetting-rate underestimates the true recognition forgetting-probability by an unknown amount. Because we see no easy way of getting quantitative estimates of the forgetting probability for the discrete models from independent recognition data, we will not use this second approach.

In order to apply the test to experimental data, we need information about the a posteriori curve characterizing the subject's recognition performance. However, many authors publish their data only in the form of ROC curves computed from the rating frequencies (see Green & Swets, 1966, for the construction of rating operating characteristics). Since the ROC and the a posteriori curve are but two ways of looking at the same data, it should come as no surprise that a property of the ROC can be derived that is analogous to the lower bound on the a posteriori curve predicted by the discrete-state models. We shall now show that the minimum slope of the ROC cannot be less than the forgetting probability, d_1 , if the discrete-state assumption holds.

The hit rate, H_k , and the false-alarm rate, F_k , corresponding to the k th point on the ROC, are found by cumulating the first k responses as "yes." Using the assumptions stated above, we find

$$\begin{aligned} H_k &= \sum_{j=1}^k P(R_j|O) \\ &= d_1 \sum_{j=1}^k r_{j1} + d_2 \sum_{j=1}^k r_{j2} + d_3 \sum_{j=1}^k r_{j3} \\ F_k &= \sum_{j=1}^k P(R_j|N) = \sum_{j=1}^k r_{j1}. \end{aligned}$$

For the slope of the straight line connecting the $k-1$ and the k th point of the ROC curve, s_k , it thus follows that

$$\begin{aligned} s_k &= \frac{H_k - H_{k-1}}{F_k - F_{k-1}} = (d_1 r_{k1} + d_2 r_{k2} + d_3 r_{k3}) / r_{k1} \\ &= d_1 + d_2 r_{k2} / r_{k1} + d_3 r_{k3} / r_{k1}. \end{aligned}$$

Since the last two terms on the right are non-negative, the following inequality holds for $k = 1$ to n :

$$s_k \geq d_1. \quad (4)$$

Thus, an alternative way of testing the discrete-state assumption is to examine the minimum slope of the ROC curve. If it goes to zero, this implies that $d_1 = 0$, that is, there must have been perfect retention according to the discrete state models. A simple graphical check is to see whether the ROC curve touches the ceiling of the unit square, that is, whether any point lies on the horizontal line connecting the points (0, 1) and (1, 1). If this occurs, the line connecting this point to the n th point of the ROC, which equals (1, 1), must have slope $s = 0$. Of course, this finding has the same implication as the a posteriori curve extending to zero.

SOME EMPIRICAL RESULTS

We have searched the recognition-memory literature for articles on experiments with yes-no designs with confidence ratings that report a posteriori curves or ROC curves for individual subjects. From the seven articles we could locate, one (Wickelgren & Norman, 1966) could not be used, because the authors eliminated all points of the ROC curves with hit rates above .98, and these, of course, are the crucial data for a test of the discrete-states assumption. All but one (Norman & Wickel-

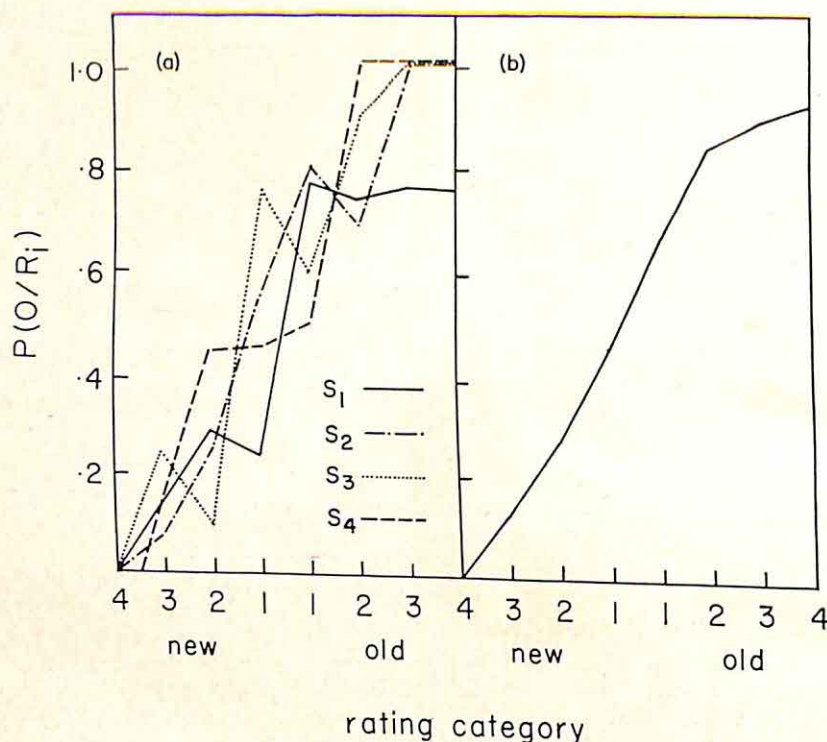


FIGURE 1. A posteriori curves obtained in Kintsch's (1968) experiment. (Section a shows the curves for the individual subjects (S_i), and Section b the curve for the combined data.)

gren, 1969) of the remaining six reports indicate some violations of the assumption.

Kintsch (1968) used a modified Peterson and Peterson (1959) procedure. On every trial, a list of five 4-consonant combinations (CCCCs) was presented at a rate of 1–1.5 sec per item, followed by a distraction task of 20-sec duration. On yes-no tests, either an item identical to one of the five list items or a new one was presented; old and new items occurred equally often, that is, $P(O) = .5$. The subjects were to give old or new responses and a confidence rating on a 4-point scale. On 60% of the trials, forced-choice tests with two, four, or eight alternatives were given. Four subjects were employed.

The analyzed data are based on trials in which items from Positions 1–4 were tested. Figure 1 presents the a posteriori curves for individual subjects and the data combined over subjects. Each curve reaches the zero level for the highest-confidence new rating. Although a statistical evaluation of the reliability cannot be given, because the number of observations on which the estimates are based is not known, it seems evident from the shape of the curves that the true asymptote

of the a posteriori probabilities is zero. Therefore, in order to account for these data, the Bernbach and the Kintsch model must assume a forgetting rate of zero.

Intuitively, the assumption of zero forgetting-probability does not seem very plausible under these conditions. It would imply that five CCCCs can be held in memory for 24 sec without any loss of information and that a single trial is sufficient to learn them perfectly. An analysis of the serial position effects on recognition performance supports this intuitive argument.

If we combine the data for yes-no and forced-choice recognition tests, the proportions of correctly recognized old items for Positions 1–5 are .71, .68, .69, .74, and .87, respectively. This serial-position effect is highly significant and, according to Kintsch (1968, p. 2), mostly due to the superior performance on items in Position 5. It follows that recognition memory for items from Positions 1–4 cannot have been perfect. Thus, the discrete-state models are rejected by these data.

Schmidt and Vorberg (Note 1) also employed the Peterson and Peterson (1959) paradigm with one highly trained subject. On

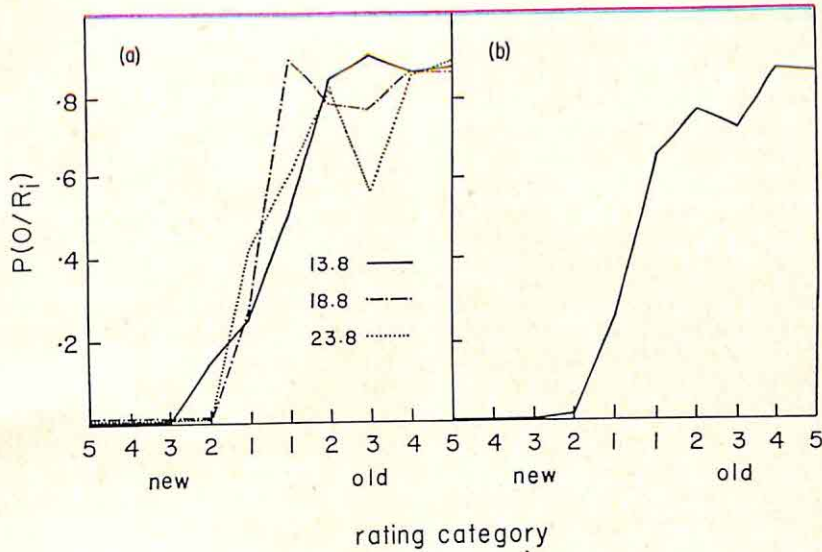


FIGURE 2. A posteriori curves obtained in Schmidt and Vorberg's (1969) experiment. (Section a shows the curves separately for the three retention intervals, 13.8, 18.8, and 23.8 sec, and Section b the curve for the combined data.)

a trial, a single 3-digit number was presented, followed by a distraction task of 13.8-, 18.8-, or 23.8-sec duration. The test was either yes-no recognition or recall. Within sessions, recognition and recall trials alternated randomly; on recognition trials, the probability of presenting an old item, $P(O)$, was .33. The subject gave his response on a 10-point scale; the 5 left-hand response categories indicated new responses. After sufficient training to reach a stable performance level, the subject was run for 36 daily sessions, with 86 trials each.

Figure 2 shows the a posteriori curves for each retention interval separately and the combined data; each curve is based on approximately 340 observations. Again, each curve asymptotes to zero. On all 585 trials on which the subject gave one of the highest three new ratings a new item had, in fact, been presented. A statistical computation shows that the hypothesis of a true a posteriori probability of .0051 or higher for these rating categories can be rejected on a significance level of at least 5%. Since a bound of .0051 corresponds to a value of d_1 of .010, we can safely conclude that according to the discrete-state models the forgetting probability must have been .010 or less.

Unfortunately, in this instance we do not have independent evidence against the zero forgetting-assumption except from recall data. The proportion of correct recalls after 13.8,

18.8, and 23.8 sec was .787, .676, and .556, respectively. Thus, there was an appreciable amount of forgetting; however, as mentioned above, this may merely reflect retrieval difficulties and does not necessarily imply nonzero recognition forgetting-rates.

Kintsch and Carlson (1967) investigated recognition learning in a paired-associates paradigm. Subjects were presented lists containing either 14, 18, or 22 pairs, with either CCCs or CVCs as stimulus members. Only two different response alternatives occurred within a given list, either 1-digit numbers, such as 1 versus 2, or letters, such as A versus E. Pairs were presented for 1.5-2 sec, a random half of them paired with the correct, the remainder with the incorrect response. Subjects were to respond *yes* or *no*, depending on whether they thought the stimulus was paired with its correct alternative or not, and to give a confidence rating on a 5-point scale. After a response, feedback was provided. Learning continued until subjects reached a criterion of two successive trials with correct response to all pairs. Four subjects were employed.

Figure 2 of the Kintsch and Carlson article (1967, p. 894) displays the rating ROC curves for recognition performance on the first trial. For three of the four subjects, the point corresponding to the laxest criterion lies on the horizontal line, (0, 1)-(1, 1), indicating a zero slope of the ROC curve and thus a forgetting

probability of zero. For the fourth subject, it comes very close. Due to the large amount of data on which the ROC curves are based, this result is very reliable. Combining the data for the three subjects with zero slope, we find that on none of the estimated 146 trials on which they used the highest-confidence new rating was the correct pair presented. This implies that we can reject the hypothesis, $d_1 \geq .020$, on a significance level of at least 5%.

What evidence exists against the assumption of zero forgetting, that is, perfect recognition memory on the first trial? Kintsch and Carlson do not report explicitly how much recognition performance improved after the first trial; however, a simple calculation shows that performance on the first trial cannot have been asymptotic. From Figure 4 of the Kintsch and Carlson (1967, p. 896) article we read a value of about .18 for the average error proportion of incorrect recognition responses on the first trial (disregarding confidence ratings). If this value were characteristic of the asymptotic recognition performance, it would take a subject an expected number of 273.0 trials to reach a criterion of two successive correct trials with a 14-item list (see Feller, 1968, p. 324). With an error rate of .18, this leads to an expected number of 49.2 errors per item, which is to be compared to an observed mean number of .73 errors per item reported by Kintsch and Carlson. Thus, it is clear that the assumption of asymptotic recognition performance on the first trial cannot be considered seriously, which implies that the forgetting probability on the first trial must have been appreciably larger than zero. Since the obtained minimum slopes of the ROC curves contradict the assumption of nonzero forgetting, the discrete-state models have to be rejected for these data.

Murdock (1974, p. 27) presents data from a single subject from a study on word recognition run by Murdock and Duffy (1972). On a trial, a list of 15 common two-syllable English words was presented at a rate of .8 sec/word, followed by a 30-word test list containing the original 15 words and 15 distractors. To each test word the subject responded on a 6-point rating scale; the 3-left-hand rating categories indicated new responses.

For the subject, the a posteriori probability for the highest-confidence new response is estimated as .036, which corresponds to a value of $d_1 = .037$, since $P(0) = .5$. Due to the large number of observations, these data reject a value of $d_1 \geq .048$ on a significance

level of $\alpha \leq .05$. Although the a posteriori curve does not go down to exactly zero, this may well have been due to the few rating categories employed, as the steepness of the curve suggests (see Figure 2.11, Murdock, 1974, p. 33). Again, in this instance we do not have independent evidence on the true forgetting rate; however, we know from comparable studies that there are huge serial-position effects that are not apparent here, because Murdock's data are averaged over list positions. For example, Shiffrin (1970) reports results from an experiment in which a list of 30 common words presented at a rate of 1 sec/word was followed by 30 test pairs; for each pair subjects had to identify the old word. The data show the typical serial position curve, with primacy and recency parts; the proportions of correctly recognized words, corrected for guessing, range from about .53 for the middle portion of the curve to .82 for the last serial position. Assuming that similar effects were present in Murdock and Duffy's (1972) experiment, the forgetting rate averaged over list positions could not have been close to zero, as implied by the a posteriori curve, thus discrediting the discrete-states assumption.

In a recent study comparing short-term retention tested by recall, yes-no, and *m*-alternative forced-choice recognition, Murdock and Wells (1974) used lists of six 2-syllable word pairs presented at a rate of either 1.5 sec or 3.0 sec per pair. After a delay of 1 sec, one of the list pairs was tested. On half of the yes-no tests, the test pair was identical to one of the pairs in the list; on the remaining half, it consisted of a stimulus member of the list paired with a new response term. If the probe was from position 5, the distractor was from position 6, and vice versa. Subjects responded on a 6-point rating scale; the three left-hand categories denoted new responses.

Murdock and Wells (1974) present the recognition data for the combined Serial Positions 1-4 and 5-6 separately for each subject and at both the slow and the fast presentation rate. For Positions 1-4, none of the individual a posteriori curves comes down to zero, although those for Subject 5 are quite close (.014 and .030 for slow and fast presentation, respectively). However, for items from Positions 5-6, five of eight for the slow condition and three of eight for the fast condition estimated a posteriori probabilities are exactly equal to zero for the highest-confidence new category. Combining over all subjects, the lowest values on the a posteriori curves are .011 and

.033, respectively, corresponding to forgetting probabilities of .011 and .034, respectively. The data are stable enough to reject forgetting rates larger than .029 and .061 for the slow and the fast condition, respectively, on a significance level of $\alpha \leq .05$.

Is it reasonable to assume that the recognition forgetting-probabilities for items from Positions 5-6 were practically zero? Murdock and Wells (1974) present some data that make that assumption highly implausible. First, after suitable corrections for guessing we find a mean correct recognition rate of .895 for these items, as estimated from the forced-choice data, averaged over subjects, presentation rates, and number of alternatives. This clearly leaves room for improvement. Second, from the recall data we see that items from Position 6 were remembered better than items from Position 5 (for slow presentation, the proportions correct for Positions 5 and 6 are .836 and .929, respectively; for fast presentation the corresponding figures are .830 and .924). Both differences are significant by *t* tests of the differences of the arc-sine transformed proportions, $t(7) = 2.82$, $t(7) = 2.26$. In view of the nearly identical average correct recall and recognition rates (.881 and .895, respectively) we tend to believe that corresponding serial-position effects must have been present in the yes-no recognition performance. Further evidence on the completely analogous effects of list positions on recall and recognition probabilities comes from a study by Bernbach and Kupchak (1972). In a comparable experiment employing lists with five paired-associates, they found the usual serial-position effects both for recall and yes-no recognition. The effects were present both in empirical indexes of retention (d' and proportion correctly recalled) and in parameter estimates of the recognition and recall probabilities of their model. Taking these arguments together, it seems safe to conclude that the assumption of zero forgetting for the average recognition performance on items in Positions 5-6 is untenable. Again the discrete-states assumption has to be rejected.

DISCUSSION

Our review of some pertinent experimental results has shown that the prediction derived from the discrete-state models of recognition memory is not supported by the data. The models proposed by Bernbach (1967) and Kintsch (1967) can account for perfect recognizability of new items that obtains in most

studies only by assuming zero forgetting-probabilities. This clearly contradicts the evidence given for the forgetting losses in these experiments. Admittedly, the reasoning by which nonzero forgetting-probabilities were inferred was indirect in some instances; however, the fact that all the data presented² are in agreement supports our conclusion that the discrete-state models are inadequate.

The experiments reported employed a variety of recognition tasks: single items presented in a Peterson and Peterson (1959) paradigm, items presented in lists, and paired-associates recognition learning. Also, diverse materials were used: 3-digit numbers, 4-consonant combinations, trigram-letter and trigram-number pairs, single words, and word pairs. Thus, the rejection of the discrete-state models of Bernbach (1967) and Kintsch (1967) is not confined to a limited class of recognition situations; rather, it seems that these models suffer from a basic shortcoming. This conclusion also holds for the recognition model proposed by Anderson and Bower (1972), to the extent that this model can be applied to the studies employing word recognition.

It would be wrong to conclude that the data presented lead to a rejection of the whole class of discrete-state models of recognition memory. The problem with the specific models of Bernbach and Kintsch is that they do not assume a separate state for new items; instead, the new-item state, D_1 , is also entered by old items unless memory is perfect. However, in order to account for perfect recognition of new items, these models have to keep new and old items in separate states, which can only be achieved by setting the forgetting probability to zero. Thus, it should be clear that discrete-state models containing a state that can only be occupied by new items but never by old ones have no difficulty in predicting perfect recognition of new items.

Discrete-state models with this property can be derived from Bower's (1967) multicomponent theory, which assumes that items are encoded by vectors of components. In one version of the theory, forgetting is seen as loss of components. At the recognition test, the subject compares the retrieved component vector of the item tested with the encoded vector of the test item. Given that there is no problem of retrieving the correct vector from memory, mismatches between the components of both vectors can only obtain if the

² Except those of Norman and Wickelgren (1969) mentioned earlier.

test item was new. Therefore, the occurrence of mismatches could be used by the subject to recognize new items correctly. Mismatches do not have to occur frequently at all for a subject to reach a posteriori values of zero. If the subject is able to confine his high-confidence new ratings to those few trials where mismatches obtain, he will achieve zero a posteriori values in spite of forgetting of components, and high similarity between the new and the old item distribution. Thus, it is apparent that perfect recognizability of new items in spite of nonzero forgetting is not at all crucial for models with a finite number of discrete states. Instead of giving up the discrete-state notion in general it might be worthwhile to consider recognition models based on Bower's (1967) multicomponent theory more seriously. However, an entirely satisfactory application of the theory to the different kinds of recognition tasks requires a more thorough investigation of the retrieval processes by which the corresponding memory traces are found.

REFERENCE NOTE

1. Schmidt, R., & Vorberg, D. *Das Kurzzeitgedächtnis beim Reproduzieren und Wiedererkennen* (Tech. Rep. 69-1). Darmstadt, West Germany: Institut für Psychologie der TH Darmstadt, 1969.

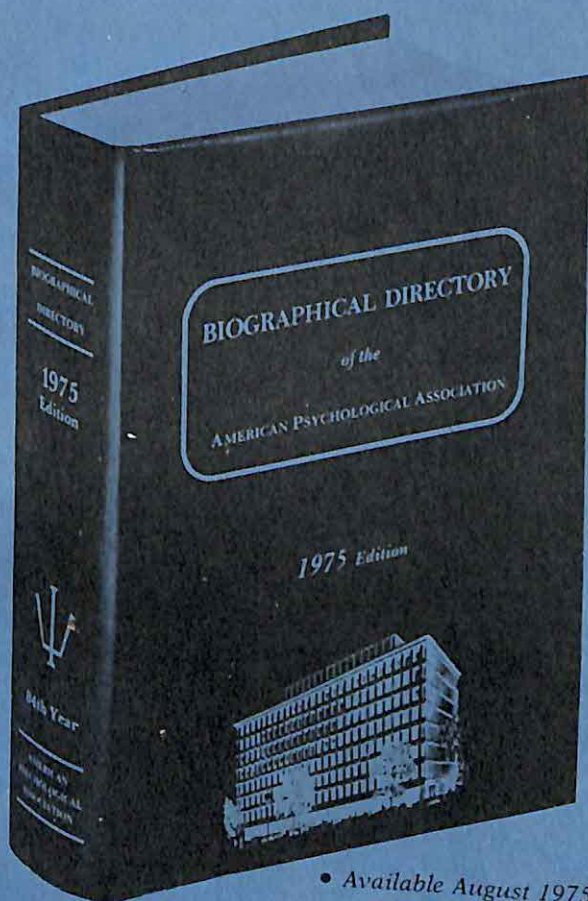
REFERENCES

- Anderson, J. R., & Bower, G. H. Recognition and retrieval processes in free recall. *Psychological Review*, 1972, 79, 97-123.
- Bernbach, H. A. Decision processes in memory. *Psychological Review*, 1967, 74, 462-480.
- Bernbach, H. A. Replication processes in human memory and learning. In G. H. Bower & J. T. Spence (Eds.), *The psychology of learning and motivation* (Vol. 3). New York: Academic Press, 1970.
- Bernbach, H. A., & Kupchak, P. G. Recognition and recall in short-term memory. *Journal of Mathematical Psychology*, 1972, 9, 237-242.
- Bower, G. H. A multicomponent theory of the memory trace. In K. W. Spence & J. T. Spence (Eds.), *The psychology of learning and motivation* (Vol. 1). New York: Academic Press, 1967.
- Feller, W. *An introduction to probability theory and its applications* (3rd ed.). New York: Wiley, 1968.
- Green, D. M., & Swets, J. A. *Signal detection theory and psychophysics*. New York: Wiley, 1966.
- Kintsch, W. Memory and decision aspects of recognition learning. *Psychological Review*, 1967, 74, 496-504.
- Kintsch, W. An experimental analysis of single stimulus tests and multiple choice tests of recognition memory. *Journal of Experimental Psychology*, 1968, 76, 1-6.
- Kintsch, W. *Learning, memory, and conceptual processes*. New York: Wiley, 1970. (a)
- Kintsch, W. Models for free recall and recognition. In D. A. Norman (Ed.), *Models of human memory*. New York: Academic Press, 1970. (b)
- Kintsch, W., & Carlson, W. J. Changes in the memory operating characteristic during recognition learning. *Journal of Verbal Learning and Verbal Behavior*, 1967, 6, 891-896.
- Krantz, D. Threshold theories of signal detection. *Psychological Review*, 1969, 76, 308-324.
- Müller, G. E. Zur Analyse der Gedächtnistätigkeit und des Vorstellungsverlaufs, III. Teil. *Zeitschrift für Psychologie, Ergänzungsband* 8, 1913.
- Murdock, B. B., Jr. The retention of individual items. *Journal of Experimental Psychology*, 1961, 62, 618-625.
- Murdock, B. B., Jr. *Human memory: Theory and data*. Potomac, Md.: Lawrence Erlbaum Associates, 1974.
- Murdock, B. B., Jr., & Dufty, P. O. Strength theory and recognition memory. *Journal of Experimental Psychology*, 1972, 94, 284-290.
- Murdock, B. B., Jr., & Wells, J. E. Parameter invariance in short-term associative memory. *Journal of Experimental Psychology*, 1974, 103, 475-488.
- Norman, D. A., & Wickelgren, W. A. Strength theory of decision rules in retrieval from short-term memory. *Journal of Mathematical Psychology*, 1969, 6, 192-208.
- Peterson, L. R., & Peterson, M. J. Short-term retention of individual verbal items. *Journal of Experimental Psychology*, 1959, 58, 193-198.
- Shiffrin, R. M. Memory search. In D. A. Norman (Ed.), *Models of human memory*. New York: Academic Press, 1970.
- Wickelgren, W. A., & Norman, D. A. Strength models and serial position in short-term recognition memory. *Journal of Mathematical Psychology*, 1966, 3, 316-347.

(Received December 14, 1973; revision received March 4, 1975)

1975 Biographical Directory

of the American Psychological Association



- Available August 1975
- 1500 pages (estimated)
- Listing approximately 39,500 APA members
- \$20 to APA members; \$40 to nonmembers and organizations.

Ordering Information: Please send your request accompanied by full payment or institutional purchase order to:

American Psychological Association
Publication Sales Dept.
1200 17th Street, N.W.
Washington, D.C. 20036



This revised and expanded **Directory** lists name and address, telephone number, education, present major field, areas of specialization, licensure or certification as a psychologist, certification as a school psychologist, ABPP and ABPH diplomate status, and APA membership status. The **Directory** also contains a geographical index section and a membership roster for each APA division. APA Bylaws, present and past APA officers, Ethical Standards, and current data on laws governing the practice of psychology in the United States and Canada are included.

1972-74 Cumulative Indexes to Psychological Abstracts

These three-year **Indexes** contain all records published in the *Psychological Abstracts* from January 1972 through December 1974. Two kinds of indexes are available:

The **Subject Index** consolidates references to the world's literature in psychology and related disciplines that were summarized in 36 monthly issues of *Psychological Abstracts*. Index terms representing the major concepts described in this literature have been applied. Phrases describing subject matter content appear under each term, enabling the user to select those references which are relevant to their needs and interests. Entries in the index contain subject index terms, *see* and *see also* reference listings, descriptive phrases, and *PA* volume and record number (two volumes).

The **Author Index** contains references to authors who produced material that was summarized in 36 monthly issues of *Psychological Abstracts*. As many as four authors are referenced for each entry. Thus, all the publications of individual investigators who conduct programmatic research in collaboration with others can be easily identified. Primary entries in the index contain up to four author's names, title of source document, year of publication, volume and issue numbers, inclusive pagination, and *PA* volume and record numbers. Secondary authors are listed by name only, with a cross-reference to the primary author and *PA* volume and record number (one volume).

The **Cumulative Indexes** for 1972-74 utilize index terms from the new *APA Thesaurus of Psychological Index Terms*. The size of the **Indexes** remains compat-

ible with those published previously. Pages are 10 x 14 inches, printed on durable acid-free paper, and encased in Class A library binding.

	U.S. and other PUAS countries	Non-PUAS countries
1972-74 Cumulative Subject Index to PA (two volumes)	\$380	\$390
1972-74 Cumulative Author Index to PA (one volume)	\$235	\$245
Both Indexes for 1972-74	\$485	\$495
1969-71 Cumulative Indexes to Psychological Abstracts are also available from APA.		
1969-71 Cumulative Subject Index to PA (two volumes)	\$360	\$390
1969-71 Cumulative Author Index to PA (one volume)	\$220	\$240
Both Indexes for 1969-71	\$465	\$495

Earlier editions of the **Cumulative Indexes**, for the period 1927-1968, may be obtained by writing to: G. K. Hall and Company, 70 Lincoln Street, Boston, Massachusetts 02111. Please do not write to APA to obtain information about these editions.

Ordering information. Please send your request for the 1972-74 and 1969-71 **Indexes** accompanied by full payment or institutional purchase order to:

American Psychological Association
Publication Sales Department
1200 17th Street, N.W.
Washington, D.C. 20036



PSYCHOLOGICAL REVIEW

GEORGE MANDLER, EDITOR

JEAN M. MANDLER, ASSOCIATE EDITOR

UNIVERSITY OF CALIFORNIA, SAN DIEGO

*AKC
Docket 2 to the
issue
librarian
2/11/75*

Vol. 82, No. 5

September 1975

Articles

A Comparison and Elaboration of Two Models of
Metacontrast Naomi Weisstein, Gregory Ozog,
and Ronald Szoc 325

Neural Stability, Sparing, and Behavioral Recovery
Following Brain Damage T. E. LeVere 344

Time-Dependent Variations in Aversively Motivated
Behaviors: Nonassociative Effects of Cholinergic and
Catecholaminergic Activity Hymie Anisman 359

Letters to the Editor 386

Editorial Consultants

The editor gratefully acknowledges the efforts of the following consultants who helped review manuscripts for this issue. They are part of a large group of psychologists who have kindly agreed to assist in the editing of manuscripts for *Psychological Review*.

IRWIN ALTMAN
NORMAN H. ANDERSON
DARVY BEM
LOIS BLOOM
PETER L. CARLTON
LOREN CHAPMAN
J. M. CHERKOFF
VIRGINIA C. CRANDALL
GERALD DAVISON
C. R. GALLISTEL
WILLIAM C. GORDON
NORMAN GRAHAM

STEPHEN A. HILLYARD
DOUGLAS HINTZMAN
DEBORAH KEMMLER
SAMUEL S. KOMORITA
DAVID T. LYKKEN
SALVATORE R. MADDI
STEPHEN A. MADIGAN
JAMES MCGAUGH
ELISSA NEWPORT
ROY W. PICKENS
D. A. RILEY
SEYMOUR ROSENBERG

EVELYN SATINOFF
ALLEN M. SCHNEIDER
DAVID J. SCHNEIDER
NORMAN SPEAR
HANS H. STRUPP
DAVIDA TELLER
W. R. UTTAL
WILSE B. WEBB
JERI D. WINE
JULIUS WISNER
ROBERT B. ZAJONC
MICHAEL D. ZEILER

The *Psychological Review* publishes articles that make theoretical contributions to any area of scientific psychology. Preference is given to papers that advance theory rather than review it, and to statements that are specifically theoretical rather than programmatic. Surveys of the literature, problems of method and design, and papers that primarily report empirical findings are ordinarily not appropriate. Letters not exceeding 250 words on topics of theoretical interest should be submitted in duplicate.

Preparation of Manuscripts: Instructions for the preparation of manuscripts appear in the *Publication Manual of the American Psychological Association* (2nd ed.). Submit three copies; one copy must be the original typed copy. All parts of the manuscript must be double spaced. In order to facilitate blind reviewing, the author's name should not appear on the manuscript or abstract, but each copy of the manuscript should include a separate title page showing the title and the author's name and affiliation. Footnotes that identify the author should be typed on separate pages. Figures should be submitted as glossy prints ready for reproduction; duplicate figures may be photocopies. Metric equivalents must be given for all nonmetric measurements. Authors should keep a copy of the manuscript to guard against loss.

Abstracts: Manuscripts must be accompanied by an abstract of 100 to 175 words typed on a separate sheet of paper. Detailed instructions for preparation of abstracts appear in the *Publication Manual* or may be obtained from the Editor.

Manuscripts: Submit manuscripts in triplicate to the Editor, GEORGE MANDLER, Department of Psychology, University of California, San Diego, La Jolla, California 92037 according to instructions provided above.

Copyright and Permissions: All rights reserved. Written permission must be obtained from the American Psychological Association for copying or reprinting text of more than 500 words, tables, or figures. Permissions are normally granted contingent upon like permission of the author, inclusion of the APA copyright notice on the first page of reproduced material, and payment of a fee of \$10 per page of APA material. Permission fees are waived for authors who wish to reproduce their own articles. Address requests for permission to the Permissions Office, American Psychological Association, 1200 Seventeenth Street, N.W., Washington, D. C. 20036.

Subscriptions: Subscriptions are available on a calendar year basis only (January through December). Nonmember rates for 1975: \$15 domestic, \$16 foreign, \$3 single issue. APA member rate: \$6. Write to Subscription Section, APA.

Reprints: Authors may order reprints of their articles from the printer when they receive galley proofs.

Back Issues and Back Volumes: For information regarding back issues or back volumes write to Order Department, APA, Box 1017, Greenwich, Connecticut 06830; University Microfilms, Ann Arbor, Michigan 48106; or Princeton Microfilms, Princeton, New Jersey 08540.

Microform Editions: For information regarding microform editions write to any of the following: Johnson Associates, Inc., P.O. Box 1017, Greenwich, Connecticut 06830; University Microfilms, Ann Arbor, Michigan 48106; or Princeton Microfilms, Princeton, New Jersey 08540.

Change of Address: Send change of address notice and a recent mailing label to the attention of the Subscription Section, APA, 30 days prior to the actual change of address. Undelivered copies resulting from address changes will not be replaced; second-class mail is not forwarded unless subscribers notify the local post office in writing that they will guarantee second-class forwarding and additional mailing offices.

ANITA DEVIVO
Executive Editor

Journal Staff

ARTHUR HAUPT
Supervising Technical Editor

KENDALL C. PATTERSON
Technical Editor

ROBERT J. HAYWARD
Advertising Manager

ANNE REDMAN
Subscription Manager

PSYCHOLOGICAL REVIEW

Copyright © 1975 by the American Psychological Association, Inc.

VOL. 82, No. 5

SEPTEMBER 1975

A Comparison and Elaboration of Two Models of Metacontrast

Naomi Weisstein
State University of New York at Buffalo

Gregory Ozog and Ronald Szoc
Loyola University, Chicago

Metacontrast has been the subject of two neural network simulations by Weisstein and Bridgeman. We compare and elaborate on the two models, correct flaws not inherent in the models' conceptualizations, and discuss the remaining shortcomings. The idea behind how U-shaped metacontrast functions are generated is similar in both models, but the assumptions about how the visual system is organized are quite different. In one of the models temporal ringing, combined with a complex and implausible linking hypothesis, is necessary in order for masking to be obtained; this model assumes a single spatial and temporal channel. In the other model, masking does not depend on temporal ringing; this model assumes multiple spatial and temporal channels and a simple linking hypothesis. We show that of the two models only the second adequately predicts empirical metacontrast functions, and we relate this model to recent evidence that the visual system contains multiple channels. Sometimes treated as a puzzling and somewhat singular phenomenon, isolated from the "mainstream" of visual data and theory, metacontrast may, on the contrary, turn out to be one of the more interesting manifestations of a multiple-channel visual system.

Metacontrast backward masking has been the subject of two neural network simulations in this journal (Bridgeman, 1971; Weisstein, 1968). Metacontrast would seem innocuous enough as a backward-masking design, and might not generate such interest were it not for the data it produces. If a

briefly flashed target (generally a disc or rectangle) and a briefly flashed mask (generally an annulus or two rectangles that flank the target) are above threshold and of roughly equal energy, the greatest reduction in apparent brightness of the target will occur when the mask is presented as much as 125 msec after the target has been turned off (Alpern, 1953). But when the target and mask are presented simultaneously, little if any reduction in apparent brightness occurs. This reduction in apparent brightness with increasing delays between target and mask is illustrated in Figure 1. It is the U in this function that generates such interest in metacontrast. Why should a target dis-

This study was supported in part by grants PHS EY 00143 (now EY 01330) from the National Eye Institute and CA 06475 from the National Cancer Institute.

We would like to thank Ronald Growney, Bruno Breitmeyer, Kevin Berbaum, and John Tangney for their helpful comments and criticisms.

Requests for reprints should be sent to Naomi Weisstein, State University of New York, Buffalo, New York 14226.

appear *only* if a mask is delayed by some 60–125 msec and not before?¹

LATE ACTIVITY AS AN EXPLANATION FOR METACONTRAST

Most theories attempting to explain the U-shaped functions in metacontrast have a common assumption, namely, that the kinds of masks used in metacontrast suppress or alter the late portion of the response to the target, and do not interfere with earlier portions. If so, in order for masking to occur, the mask must be delayed until its presentation will interfere with the appropriate (late) portion of the target response.

The response to the target may be divided into early and late portions if for no other reason than that a brief stimulus flash produces neural activity which may last well over 150–200 msec after the stimulus goes off (Sturr & Battersby, 1966; measurement of the gross potential).² In addition, the

¹ One way around this puzzle is to assume that U-shaped functions in metacontrast do not really exist (e.g., Eriksen, Becker, & Hoffman, 1970; Lefton, 1973a). However, proponents of this position will admit that there is a U-shaped reduction in apparent brightness of the target (Eriksen et al., 1970; Lefton, 1973b); their main criticism is that *detection* functions in metacontrast are not U-shaped. This is generally true (but see Cox & Dember, 1972); however, it is an irrelevant argument. There are a number of similar paradigms which demonstrate that detection and apparent brightness do not covary. Cornsweet and Teller (1965) found that increment thresholds remained constant even though the apparent brightness of a steadily presented disc varied depending on the brightness of a surrounding annulus. Few would claim that although the increment threshold to the disc did not change, simultaneous brightness contrast is not "real": discs do look brighter when surrounded by dark annuli than when surrounded by bright annuli. One simply has to use a response indicator appropriate to the phenomenon one wants to measure.

Another approach is to assume a priori that complex phenomena such as metacontrast are not worthy of serious attention (e.g., Sekuler, 1973). However, it would seem that we progress by exploring, rather than evading, that which is puzzling.

² As early as 1966, Sturr and Battersby noticed the possible relevance of the later portions of neural activity to backward masking. They hypothesized that the entire 150–200 msec of neural activity had to occur undisturbed in order for a

early portion of this activity looks different from the later portions. If early and late correspond to different perceptual events, then the late-activity explanation for metacontrast may make sense. In metacontrast, something about the *appearance* of a target is affected, while its detectability is generally unchanged: Its apparent brightness or contrast decreases, and its contours break up or fragment. The loss in detectability is held to a minimum by the metacontrast design: Both target and mask are well above threshold, and they are adjacent to each other rather than superimposed on the same retinal location. If we suppose appearance to be connected with the late portion of neural activity, then a mask, to be effective in metacontrast, might have to coincide with the late portions of neural activity to the target. Hence U-shaped masking functions would be predicted. In somewhat loose terms, then, two adjacent, briefly flashed, above-threshold stimuli might not be able to affect each other's appearance unless the activity produced by one of the stimuli "hits" the right portion of the activity produced by the other stimulus; if this portion of activity is late activity, then U-shaped functions will be generated.

NEURAL MODELS AND VISUAL-SYSTEM ORGANIZATION

To tie these ideas down to specific predictions, one may make a model of neural response or visual-system activity. This model should generate metacontrast functions, be consistent with what is known about mammalian visual-system organization, and make reasonable *linking hypotheses*—that is, make reasonable hypotheses about how neural response might correspond to the perceptual data (Brindley, 1970).

Before considering each metacontrast simulation in particular, a few comments are in order about two somewhat competing views of general visual-system organization, the

stimulus to escape masking. This is a somewhat more general form of the theories developed to account for metacontrast; additional assumptions, about what happens when, have to be made in order to predict U shapes.

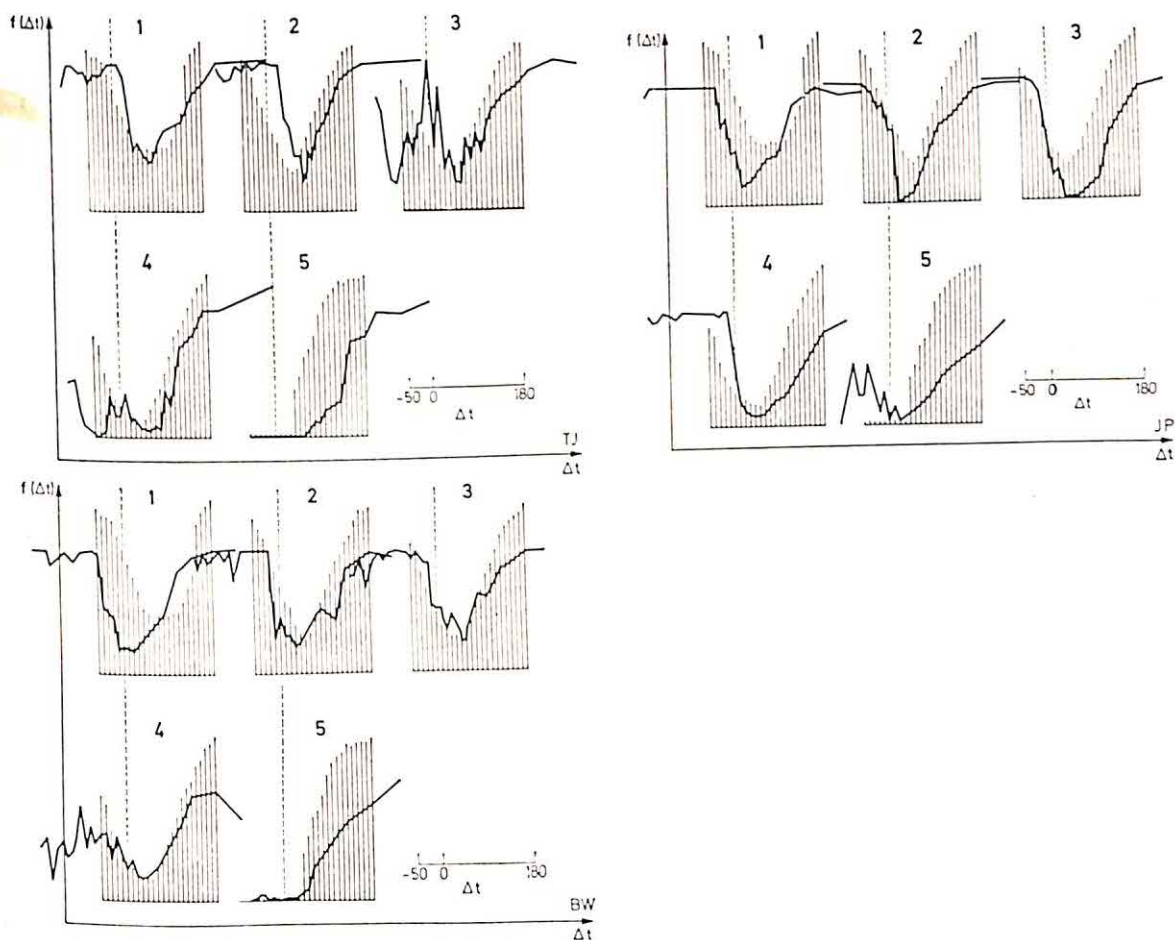


FIGURE 1. Metacontrast functions for each of three subjects, TJ, BW, and JP, as indicated on the figure (solid curves) at each of five target-to-mask energy ratios, T/M (functions 1 through 5; from Weisstein, 1972). The mask was always a lighted 16 msec, 16 footlambert (54.8 cd/m^2) annulus; the target was a lighted, 16 msec disc whose luminance decreased respectively from 1 through 5 as follows: 16, 8, 3.2, 2, and 1 footlambert ($54.8, 27.4, 10.9, 6.8, 3.4 \text{ cd/m}^2$). Each point on the solid curves represents a geometric mean of five magnitude estimations of the apparent brightness of the disc at each of 27 Δt s, beginning with $\Delta t = -100$ msec and proceeding in 10 msec increments to $\Delta t = 100$ msec; thereafter in 25 msec increments to $\Delta t = 200$ msec, and finally, 250 and 300 msec. U shapes are found for each subject for each T/M except the lowest one or two. Dotted vertical lines indicate simultaneous presentation of target and mask, or $\Delta t = 0$. If the function falls below the amplitude at the dotted vertical line, then it can be said to be U-shaped. The height of the solid vertical lines represents the points predicated by Weisstein's (1972) model. (From "Metacontrast" by Naomi Weisstein, in D. Jameson & L. Hurvich (Eds.), *Handbook of Sensory Physiology* (Vol. 7, part 4: *Visual Psychophysics*). Copyright 1972 by Springer-Verlag. Reprinted by permission.)

single-channel view and the multiple-channel view.

Single Channel

A single-channel view assumes that visual-system response can be characterized by one set of parameters. To put it in terms of neurons, a single spatial and temporal chan-

nel (or, equivalently, filter) would correspond to the assumption that the visual system responded as if all single units had the same receptive field size and time course of excitation and inhibition. Then one would simply need one function to describe the visual system's spatial properties and one function to describe the visual system's tem-

poral properties. The single-channel assumption is not necessarily contradicted by the evidence that single units in the visual system have different receptive field sizes and different rates and latencies of firing, since all units could contribute to a composite size and response rate. (See Nachmias, 1968, for a discussion of a composite visual-system spatial response, and see Cornsweet, 1970, for an excellent discussion of how if one assumes a single channel, one can obtain an overall estimate of the visual system's temporal and spatial response by utilizing frequency-response techniques.)

Multiple Channel

A multiple-channel view assumes that to characterize visual-system response, families of parameters are necessary. Many different channels may be active in the visual system, depending on whether a stimulus is small or large, periodic or aperiodic, transient or sustained. These channels do not contribute to one composite, overall function; rather, at some stages in the visual system their activities are independent of each other. Thus, for instance, at threshold it would not be the sum of the activity of a number of different channels that determines threshold, but the activity of a particular channel that has reached threshold, activity not directly influenced by the subthreshold mumblings of other channels that may be active. (The threshold itself may be lowered somewhat by subthreshold mumbling due to probability summation, but this would reflect the independent activities of a number of channels rather than the change in activity of a particular channel as a result of the activity in another. See Graham, *in press*, for an excellent discussion.)

An above-threshold view of a multiple-channel model might assume that different channels are responsible for different features of a stimulus—some channels for the edges of the stimulus, for example, and others for the inside. We will have more to say about this later. In terms of neurons, a spatial multiple-channel model would assume that units with different receptive field sizes respond to stimuli that are somewhere

within the range of their field size, and that when a stimulus is too small or too large, the unit will simply not respond. (Spatial multiple-channel models have usually been phrased in terms of parallel spatial filters, each responding to a band of spatial frequencies, but we prefer, for the moment, to use a somewhat looser notion.) A temporal multiple-channel model would assume that units would respond at different rates, and this implies that some units might prefer transient stimulation and other units might prefer sustained stimulation. Again, we will have more to say about this later. The multiple-channel model doesn't mean that only one channel will respond if only one stimulus is presented; a particular stimulus may produce response in a number of channels. The important point is that these channels are looking at different aspects of the stimulus, and they are responding at different rates.

THE TWO METACONTRAST SIMULATIONS

Of the two attempts to quantify the hypothesis that metaccontrast is the suppression of the late activity or slow portion of visual-system response to a stimulus, one (Weinstein, 1968, 1972) is a multiple-channel model, and the other (Bridgeman, 1971) is a single-channel model. We will examine the two models to see which one seems the more plausible. Neither model is adequate as it stands, and in considering which model better fits the data, we need to expand and elaborate on both models. First, we will present Bridgeman's simulation, which assumes a single channel, taking the steps necessary for it to generate metaccontrast functions (this was not done in his paper). As we examine this model, some serious shortcomings that have previously escaped notice will become apparent. Then we will consider Weinstein's simulation (1968, 1972), which assumes multiple channels, expanding it so that it can deal with spatial interactions. We will show that two shortcomings noted for this model as it was formulated may be more apparent than real. Finally, we will present an elaboration of this model, which associates the fast and

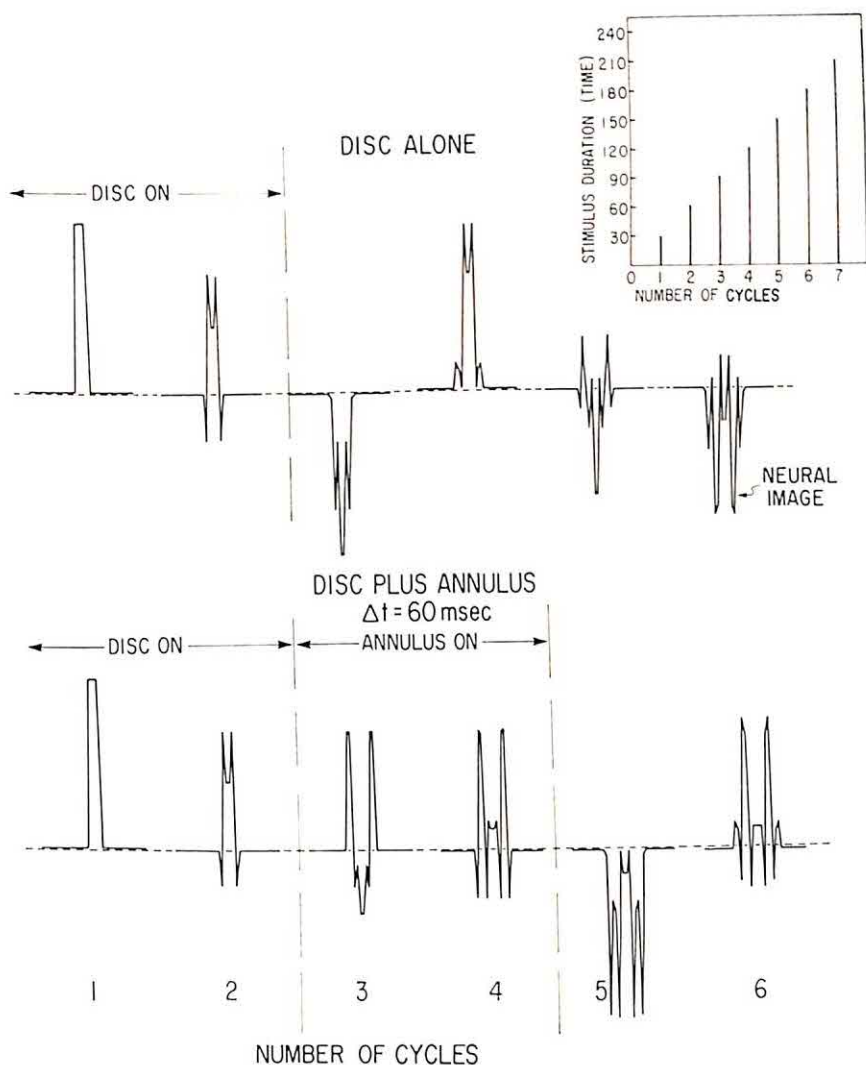


FIGURE 2. The activity of Bridgeman's (1971) neural network (the neural images) for each of six successive time periods. The top row of Figure 2 shows neural images that occur during and after a disc is presented. The bottom row of Figure 2 shows neural images that occur during and after a disc is presented, followed by an annulus. The main thing to notice is that the neural images of cycles 3 through 6 for disc alone (top row) versus disc plus annulus (bottom row) do not resemble each other: This is the basis of the metacontrast interaction. Time is a discrete function of number of cycles: This is plotted in the inset.

slow responses assumed by Weisstein with the recent work on transient and sustained mechanisms (channels) in the visual system.

Metacontrast Simulation 1

Bridgeman (1971) hypothesized that metacontrast occurs because the late neural activity to a target followed by a mask differs from what it would be to a target alone.

To obtain a picture of neural activity, Bridgeman uses the one-dimensional neural network developed by Ratliff and his co-workers (1965), in which each neuron inhibits and is inhibited by its neighbors at recurrent intervals. A neuron in the system inhibits its neighbors in proportion to its own activity and to its distance from its neighbors. In addition, inhibition is subject to time delays: Adjacent neurons are inhibited with a

time lag of 30 msec and neighbors one removed are inhibited with a time lag of 60 msec. To simulate this digitally, Bridgeman uses the Ratliff equations (1965, p. 110), with parameters chosen from Barlow's (1969) study of *Limulus*. This network is single channel because it has a single set of spatial and temporal parameters—a spatial weighting function consisting of a point excitation and a spread of inhibition on either side, and a temporal weighting function that is simply a step for excitation and inhibition from adjacent units, and a step with a one-period lag for inhibition from units one removed.

The Network

The Ratliff (1965) equations represent a discrete form of repeated inhibition: Each recurrence of inhibition "spreads" a signal laterally so that one by one, units further and further removed from the site of stimulation become involved. The activity of the neurons in the network, what we may call the *neural image*, is shown in Figure 2, during and after stimulation, for each successive recurrence of inhibition of the network. The top row shows a network responding to a target alone, and the bottom row shows a network responding to a target and then a mask, which is presented after the initial response and one cycle of inhibition of the network. The inhibition delay defines time in this network; that is, one cycle means that up to 30 msec have passed, two cycles means that up to 60 msec have passed, and so forth. This quantization or "stepping" of time is shown in the inset of Figure 2.³

³ This figure was prepared using the output from our own simulation program, which repeats the calculations described in Bridgeman (1971). The figure agrees with Bridgeman's Figure 3 (p. 533) with two exceptions (see below), and for the sake of brevity we have left out his third row, which shows the network response to the annulus alone. Our Figure 2 also differs from Bridgeman's because we adjusted the neural images so that they each had the same maximum peak-to-trough amplitude; this way, it is easier to see what the neural images look like for the later cycles where the total signal is very small. (Adjusting the peak-to-trough amplitude, or normalizing these

As shown in Figure 2, when the stimulus is on, the corresponding neural images remain fairly similar to the physical stimulus through successive cycles, that is, through successive 30 msec intervals at which the inhibition recurs. Each successive recurrence can be seen to spread the signal, and edges and troughs develop at the borders of the stimuli. (Edges and troughs generated by networks like these are the basis of many explanations of Mach bands and related effects.) Bridgeman adopts the term *homophotic* from Stigler (1910) to describe the set of neural images during stimulation. The term *metaphotic* describes the set of neural images that occur when the stimulus is removed. Although the metaphotic images no longer resemble the physical stimulus, they are uniquely related to it.

Masking

The set of neural images that occurs when the target is presented by itself comprises the reference to which all other conditions are compared. The amount of masking is assumed to depend on the extent to which the set of neural images formed when both target and mask are presented differs from the set of neural images formed when the target is presented by itself. This measure

functions, does not change the measure of similarity that Bridgeman used and we have replicated, since the measure itself normalizes all functions.) While this effectively makes all the neural images the same size, it should be stressed that by the seventh or eighth cycle, the actual neural image is almost indistinguishable from noise.

There are some minor differences between our top-row cycles (3 and 4), and Bridgeman's. Below, there are some minor differences between our similarity function (Figure 3), and Bridgeman's. We have gone through extensive checking procedures on our own simulation (following Bridgeman's description of how he did his simulation) and are convinced that our calculations are correct. The differences may be due to the following factors: We used different computers, which could have led to a difference in word size, number of significant digits, rounding errors, plotter and plotting routine; and there may have been some small drafting errors in Bridgeman's figures. In any case, none of these differences are sufficiently large to affect the arguments that we make in this paper.

of similarity includes the entire neural image, from one end of the network to the other: The activity of *every* element in the network is considered, rather than, for instance, just the activity in and around the target area. We shall have more to say about this below. The similarity comparison is made cycle by cycle, and a similarity function is produced, which gives a value of similarity for pairs of neural images (target, target + mask) for each cycle.⁴

Bridgeman (1971) divides masking into two types: The type associated with the homophotic neural images (forward and simultaneous masking) and the type associated with the metaphotic neural images (backward masking and metacontrast). According to Bridgeman, homophotic neural images allow subjects to make *detection* decisions, while metaphotic neural images have to do with the *appearance* of the target. It is not clear how the visual system is able to distinguish homophotic from metaphotic images; the images themselves cannot provide a unique criterion for whether or not a stimulus is physically present. But for the purposes of argument, let us assume that this problem can be solved. If we assume, then, that the system is somehow able to tell the difference between homophotic and metaphotic, in those instances where

the mask is presented before or simultaneously with the target, the homophotic images of the target are somewhat different from those of the target and mask together. In this case, detection drops. When the mask is presented later than the target, the homophotic images are identical, but the metaphotic images of the target alone are different from those of the target-plus-mask, as can be seen in Figure 2. Here, then, the appearance of the target—its apparent brightness—is assumed to change, and one has the makings of a metacontrast situation. (Actually, the model predicts that whenever detection drops, the appearance of the target also changes for those instances where the target is detected, since if the homophotic neural images of target and target-plus-mask are dissimilar, so are the metaphotic images. But the important point is that one can have appearance changes without detection changes, and this forms the basis of the metacontrast predictions.)

Thus the appearance of a stimulus is mediated by the late neural images (i.e., the late neural activity), and metacontrast occurs when these late neural images differ from what they would be if a target had been presented alone. How is a metacontrast function constructed from this formulation?

Constructing the Metacontrast Function

A metacontrast function is a plot of some measure of the change in appearance of the target (generally apparent brightness) for each delay between target and mask (for each Δt). No plots of this kind appear in Bridgeman's (1971) simulation. Rather, neural image similarity is computed cycle by cycle for one Δt in the backward masking range, one in the forward masking range, and one at simultaneous masking. The similarity function obtained is different for these three cases; in particular, for the backward masking case the later, metaphotic neural images of the target do not resemble the corresponding images of the target-plus-mask. This similarity function is shown in Figure 3.

A similarity function is not a metacontrast function, because it provides informa-

⁴ These similarity comparisons are called cross-correlations, but the reference appears to be to a Pearson r . That is, if one computes the actual cross-correlations for the output generated by Bridgeman's (1971) network for the 6 cycles illustrated in Figure 2, they are, in order, 1600, 256, -394, 172, 16, and -5. This bears no resemblance to Bridgeman's Figure 6. Next we tried various variations on our cross-correlations, but none of them got us near Bridgeman's Figure 6. We thought perhaps he had normalized the cross-correlations in some way, so we computed a "constant of correlation" (Lathi, 1965): $F(x) * G(x)/F(x) * F(x)$, where $*$ denotes convolution (for symmetrical stimuli, this is the same as cross-correlation), $F(x)$ = the neural image in response to the target alone, and $G(x)$ = the neural image in response to the target-plus-mask. Again, these numbers in no way resembled Bridgeman's Figure 6. We then computed Pearson rs [the normalized cross-correlation for the points $F(0)$, $G(0)$], and finally our results agree in all major respects with Bridgeman's Figure 6 (for the minor respects in which they do not agree, see footnote 3).

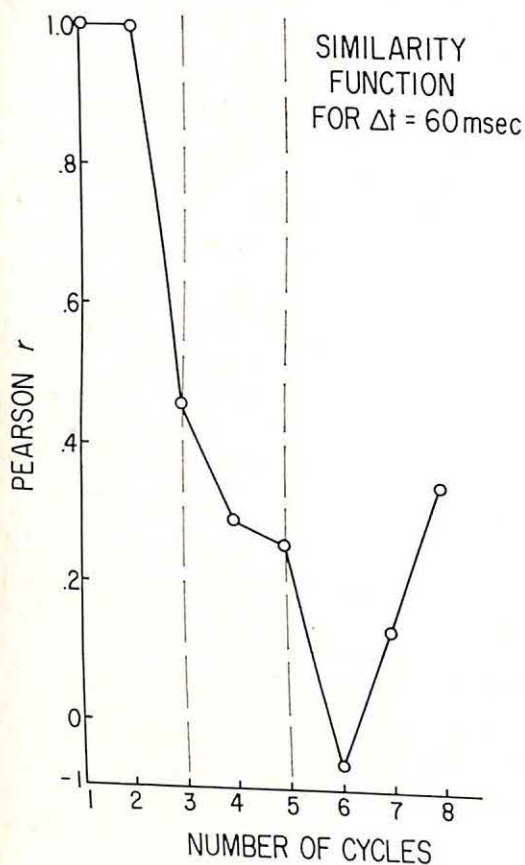


FIGURE 3. The similarity function (what Bridge-man, 1971, calls the cross-correlations) for what is assumed to be the metacontrast or backward masking case (actually, this represents only the point $\Delta t = 60$ msec). The points on this function are Pearson r s for the neural image pairs (disc alone, disc-plus-annulus) which occur during the same cycle. Each column in the top and bottom row (each cycle) in Figure 2 allows calculation of another Pearson r .

tion about only one delay (or value of Δt). But the information from the similarity function can be suitably processed to obtain a metacontrast function. We can decide on an appropriate measure for masking (a measure that computes from the similarity function for a particular delay a number that corresponds to target apparent brightness for that delay) and then we can repeat this calculation on similarity functions for a sufficient number of delays to give us a metacontrast function.

If we take these calculations step by step, we must first decide on what information

to compute from the similarity function. The computation will yield a number, which will be assumed to correspond to target apparent brightness. One should remember that the model assumes that the apparent brightness of a target is mediated by its metaphotic images, and that masking is related to how similar the neural images that occur when a target is flashed in conjunction with a mask are to the neural images that occur when a target is flashed alone. However, the set of metaphotic images cannot be a necessary condition for target apparent brightness, since a steadily presented disc of light has a distinct apparent brightness, and therefore homophotic stimulation is sufficient for the perception of apparent brightness. (This is related to the problem of how the system is able to distinguish homophotic from metaphotic stimulation in any case.) So a suitable measure of apparent brightness for a target for a particular Δt might consider both homophotic and metaphotic neural images, and simply sum the similarity values, that is, sum the points computed for the similarity functions for that Δt . The sum would run from the onset of stimulation (the first cycle) through the seventh or eighth cycle (the just-distinguishable signal).⁵ The greater the similarity, the greater the value of the sum, the brighter the target, and hence the less the masking.

The results of these calculations are shown in Figure 4. Figure 4a shows calculations for three points in the original network; in Figure 4b, we calculated all the other $+\Delta t$ s that can be generated by the network. These are 30, 90, 120, 150, 180, and 210 msec. (The entire set of $+\Delta t$ s is limited to about eight, due to the quantization of time and the fact that, at Δt s longer than 210, the target signal can not be distinguished from

⁵ It should be noted that we also tried summing various other combinations of cycles: three through seven (i.e., just the metaphotic cycles), four through six, and so forth. None of these sums changes the argument. However, one should not restrict measurement to the later cycles, and indeed, measurement should stop at about the seventh or eighth cycle, since after this even though the Pearson r may still take on large negative or positive values, the actual signal is indistinguishable from noise.

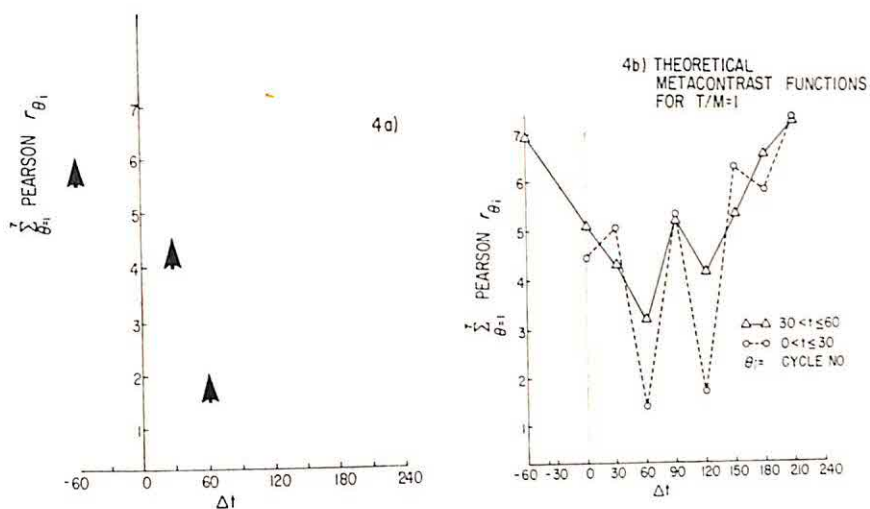


FIGURE 4. Masking, as computed from Bridgeman's (1971) similarity functions. Each point represents the sum of Pearson r s at a particular Δt from the first through the seventh cycle. Figure 4a shows what one would obtain if attention were restricted simply to the three values of Δt that Bridgeman considered; Figure 4b shows the masking functions one obtains when one computes the sum of points on each similarity function for all the other Δt s that can be generated by Bridgeman's network. The dotted line shows the masking function obtained when both target and mask are presented for one cycle of the network; the solid line shows the masking function obtained when both target and mask are presented for two cycles of the network (allowing the network one application of inhibition before stimulation is removed). These can be thought of, respectively, as a target and mask presented for durations of $0 < t \leq 30$, and $30 < t \leq 60$, where $t =$ duration.

noise.) In addition, we calculated the same set of Δt s for another duration, namely, a duration less than or equal to 30 msec, or the condition where the network does not have a chance to initiate a second cycle before the stimulus is turned off.

The arrows in Figure 4a suggest the beginnings of a U. But when one extends these calculations to include the complete set of Δt s possible in this network, it becomes apparent that the function reverses itself two or three times (rather than just once, to produce the U) and thus does not resemble metacontrast functions which are empirically obtained. Let us examine this problem, and related temporal problems, in a little more detail.

Shape of the Metacontrast Function

Networks such as the one in Figure 2, with inhibitory feedback occurring after

some short, discrete delay, are subject to damped oscillations or temporal "ringing." That is, a point on the neural image will, with sustained stimulation, first fire at a certain rate; on the next cycle, it will fire less; on the succeeding cycle, it will fire somewhat more again; and so on. This ringing allows strong masking to develop. Pairs of neural images (target, target-plus-mask) at particular cycles can become quite dissimilar, with corresponding points on the neural image even having differing signs (see Figures 2 and 4, cycle 6 for instance: The Pearson r for this cycle goes negative). But ringing also results in the metacontrast function itself oscillating a number of times. If, for a given Δt , the neural images produced by a target-plus-mask resemble the neural images produced by a target alone, then for the next Δt , the neural images of the target-plus-mask will be out of phase with those

of the target alone, and hence metacontrast masking at successive Δt s will first be weak, then strong, then weak again, then strong again, and so on. But metacontrast functions do not oscillate (Alpern, 1953; Growney & Weisstein, 1972; Schiller & Smith, 1966; Weisstein, 1971, 1972; Weisstein, Jurkens, & Onderisin, 1970). It is unlikely that if metacontrast functions had two or three major troughs and peaks they would have been missed by the sampling of Δt s utilized by the various studies. For instance, Figure 1 shows empirical metacontrast functions (solid lines) which contain 27 Δt s, 21 of them spaced in 10 msec increments, from -100 to $+100$ Δt . In the design from which these data are drawn, all the Δt s are presented in random order five times per Δt per target luminance per subject. While there are some ragged edges in Subject TJ's functions, there is no evidence of the kind of abrupt reversals shown in Figure 4b.

Other Temporal Properties of the Network

What matters for a metacontrast function is the *ratio* of target duration to mask duration (with luminance held constant), which we will call T/M . When the durations of target and mask are equal, that is, when $T/M = 1$, the same U-shaped metacontrast function is obtained, no matter what the durations of target and mask, that is, no matter what T and M (Kahneman, 1967; Weisstein, 1972; Weisstein & Growney, 1969). Thus a target and mask can be presented for 25 msec each or for 125 msec each, and the difference in the metacontrast functions is negligible (Kahneman, 1967). On the other hand, as $T/M \rightarrow 0$, metacontrast functions change shape (they become monotonic, increasing with increase in Δt , as in Figure 1). Neither of these properties of metacontrast is generated by Bridgeman's (1971) network.

Target-to-mask duration ratio approaches zero ($T/M \rightarrow 0$). Time is a discrete function of number of cycles in this network. While there is nothing wrong with approximating time by discrete steps, the assumed

duration of each step is much too large to simulate the pertinent features of metacontrast. This is especially true for the first 30 msec, when T/M is the most sensitive to variations in duration. For instance, if one presents a target for 1 msec and a mask for 16 msec, a monotonic function of $+\Delta t$ will be obtained; if one presents a target for 16 msec and a mask for 16 msec, a U-shaped function of $+\Delta t$ will be obtained (Weisstein, 1971). This network cannot distinguish between these two cases.

Target and mask of equal duration ($T/M = 1$). At the jump points of the network—30 msec, 60 msec, 90 msec—the neural images change. This means that if one holds T/M constant at 1, but multiples of 30 msec are added to both target and mask (that is, one changes the number of cycles the network goes through while each stimulus is present), different metacontrast functions will be obtained. Figure 4b illustrates this: The solid line shows a masking function obtained by allowing the network to receive the target (30 msec) and then recycle once while the target is still present, for a total target duration of $30 < t \leq 60$. The dotted line shows the masking function obtained by allowing the network no recycling while the target is on: It simply receives the target ($0 < t \leq 30$). In each case, the mask is presented for the same duration intervals as the target. Since $T/M = 1$, the masking functions should be identical. However, as Figure 4b shows, the functions are quite disparate. This contradicts a considerable body of empirical evidence (Weisstein, 1972).

These temporal problems are not caused by the fact that time is quantized in this network, but rather by the dependence of the recurrence of inhibition on the stepping of time. Hence, the problems are not solved by changing the cycle time, that is, by decreasing or increasing the duration between steps. Making the cycle time longer will produce masking at Δt s which are too great to correspond to what has been found empirically, and it will only increase the difficulties in predicting monotonic functions as $T/M \rightarrow 0$. Making the cycle time shorter

will only translate the metacontrast functions leftward and increase oscillation frequency. This can be illustrated in Figure 4b: Shortening the cycle times, for example, would be equivalent to rescaling the x -axis. If this were done, and each cycle time were assumed 5 msec long (so that the intervals were $-10, -5, 0, +5, +10$, and so on) it would solve neither the problem of oscillation nor the problem of differing metacontrast functions for $T/M = 1$ for different target-mask durations. But shortening the cycle time would introduce additional problems: Masking would occur too soon, and, at the delay at which masking *should* occur, the network would have already gone through so many cycles that the signal would be too damped to be distinguished from noise.

Spatial Properties

The network exhibits spatial as well as temporal ringing. This means that it does not predict the spatial properties of metacontrast either. Empirically, as the separation between target and mask increases, metacontrast decreases monotonically. The network, on the other hand, produces abrupt reversals in masking as the separation between target and mask is increased, unit by unit. So, for instance, masking is *greater* at spatial separations of three neural units between target and mask than it is at separation of two units.⁶ These abrupt reversals are again a function of the ringing of the network, especially at the edges of the target and mask; when these neural images are in phase, masking is slight; when they are out of phase, masking is heavy.

The Linking Hypothesis: Correlation as an Index of Apparent Brightness

Even if the temporal and spatial properties of metacontrast were better predicted by

⁶ The sum of Pearson r s for $\Delta t = 60$ msec (Bridgeman's condition, 1971) for seven cycles at a separation of two units between target and mask is 3.75; at a separation of three units between target and mask it is 2.94. If one inspects the metacontrast function for this condition (solid lines in Figure 4b), it is clear that relative to the rest of the points, this is a large difference.

this network, the way neural image similarity is measured cannot be justified. For each cycle, a Pearson r is computed for the neural images obtained in response to the target alone versus those obtained in response to the target-plus-mask. What this means is that the value of each element in the network for a target neural image at a particular cycle is multiplied by the same element in the network for a target-plus-mask neural image at that cycle, and then the products are summed across the entire network and normalized. Consequently, the measure will produce masking no matter what is added to the target, as long as it is presented somewhere else in the network (and, at least until the cycles spread the neural image to the location in the network where the additional stimulus is located, it doesn't matter *where* else in the network the additional stimulus is presented). That is, for this measure, the magnitude of masking is not affected by the separation between target and mask until the signal is spread far enough. Thus, within some finite number of cycles one could simply add a second stimulus at the very edge of the network, or some place in the middle, and the same amount of masking would be produced. Indeed, one could add two, three, many stimuli at the edges of the network and produce lower and lower correlations.

This does not agree with the evidence. Adding flanks in metacontrast only decreases the apparent brightness of a target when the flanks are close to it (Alpern, 1953; Weisstein & Growney, 1969). As the spatial separation between target and mask increases, masking declines sharply. Furthermore, when disc and annulus are presented simultaneously at the same luminance and duration, enhancement may occur (Matthews, 1974).

It might be more reasonable to inspect the change in neural activity only for those units in and around the location of the target.⁷

⁷ It could be argued that the information about the target might not be restricted to a retinally localized area. Such distributed encoding schemes are discussed by Weisstein (in press). Distribution of target information over an extensive retinal area may be a possibility, but then, generally,

When this is done, the correlations remain near 1.00 for both the simultaneous condition and for 60 msec Δt . Cycle by cycle, for 60 msec Δt , they are 1.0, 1.0, 1.0, .97, .932, .997, .999. This hardly constitutes a reduction in similarity.

Summary of Metacontrast Simulation 1

This model fails to simulate a metacontrast function, and it predicts neither the spatial nor the temporal properties of metacontrast. In addition, masking is predicted only by the use of an implausible linking hypothesis.

Many of these flaws have specifically to do with the way in which this particular simulation is constructed and the assumptions behind it. But some of the difficulties might beset any single-channel network, which essentially repeats the physical stimulus (with some modification). Such a network preserves information, but it can hardly be thought of as processing information. That is, different features of a stimulus, which might provide some basis for deciding what kind of stimulus has been presented and when that stimulus has been presented, are not independently represented, at least at the level where the network is still single channel.

We can use Bridgeman's (1971) network as an example of the problems this may create. In Bridgeman's network, apparent brightness is distinguished from detection by a stimulus-dependent event—the stimulus is turned off. But this assumption is untenable for two reasons. First, with a single channel alone, it is not possible to distinguish which is the homophotic and which is the metaphotic signal. Second, the assumption that metaphotic activity results in apparent brightness contradicts the obvious fact that if a stimulus is left on long enough for a subject to make a response to it, the subject will respond that it has a distinct apparent brightness. However, making things stimulus-dependent may be one of the few ways one can distinguish between differ-

ent stimulus properties while preserving the notion of a single channel.

It is entirely possible that a more convincing single-channel model can be constructed that will overcome some of these problems. Meanwhile, however, it might be worthwhile to turn attention to the other metacontrast simulation (which assumes multiple channels and a rather straightforward linking hypothesis) and see if this simulation does any better.

Metacontrast Simulation 2

The other metacontrast simulation (Weinstein, 1968, 1972) assumes that there is mutual inhibition between fast- and slow-responding neural populations. In particular, for metacontrast it assumes that units whose response is fast inhibit units whose response is slower. Hence, one must delay the stimulus for the fast-responding units in order to have them interfere with a prior stimulus.

Quantification and Criticisms

These assumptions were worked out quantitatively in a neural network incorporating Rashevsky-Landahl two-factor neurons—neurons whose temporal response is a combination of excitation and inhibition (Weinstein, 1968, 1972). This network successfully predicted a number of properties of metacontrast, such as (a) the shape of the metacontrast function and its change of shape as T/M changes (as in Figure 1); (b) the reappearance of the target when target and mask are repeatedly presented at appropriate Δt s (as in Schiller & Smith, 1966); and (c) for 34 out of 35 data points that appeared in the metacontrast literature, the Δt at which the minimum of the U occurred. This model is continuous, so any combination of target and mask luminance, duration, and Δt can be simulated.⁸ It does

⁸ Inspection of the predictions from Weinstein's (1968, 1972) model in Figure 1 might, at first glance, seem to indicate that the same type of quantization of time applies to her model as well. But this is not the case. Weinstein's network was simulated on an analog computer, and within the resolution of the computer, any duration and Δt can be simulated.

more than one channel is assumed, and one must use a more plausible linking hypothesis to relate such distributed encoding to perceptual events.

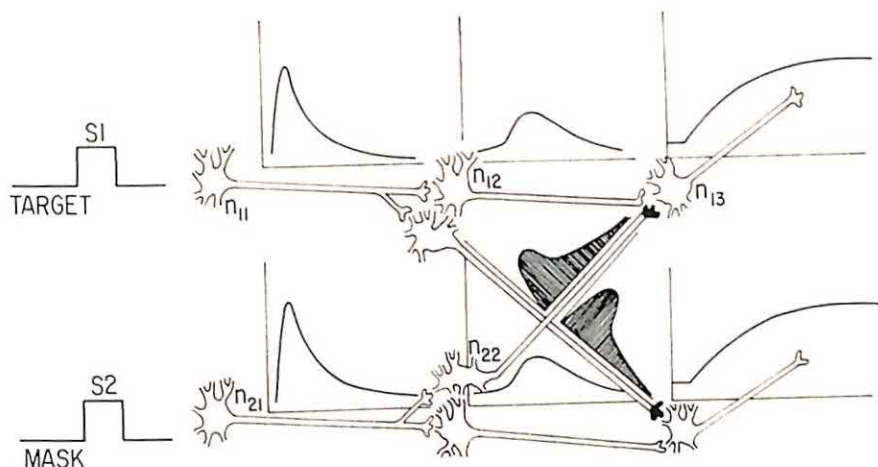


FIGURE 5. Expanded version of Weisstein's (1968, 1972) model, which clarifies the spatial symmetry implicit in that model. The subscripted neurons are those which also appear in Weisstein (1968, 1972). Neuron n_{22} responds faster than n_{12} ; therefore, stimulation for n_{22} must arrive later than stimulation for n_{12} in order that the response of n_{12} be suppressed. If one traces responses in the unlabeled neurons, it is clear that an identical effect will occur for the bottom neuron similar to n_{13} if the stimulus in the top row is delayed. Thus one can interchange target and mask and produce the same result.

not depend on temporal ringing to produce masking, and thus it is not subject to the same temporal difficulties as the network considered earlier. In addition, the measure of target apparent brightness is proportional to the frequency of firing of the unit designated as carrying information about the target, so the linking hypothesis is not quite so problematic as it is in Bridgeman's (1971) network.

However, there are two criticisms of this simulation. One has to do with what are assumed to be the spatial constraints imposed by the model, and the other with a temporal assumption behind some of the parameters chosen for individual neurons.

Spatial Interaction

The criticism here has been that "the simulation will not work if target and mask are interchanged" (Bridgeman, 1971, p. 530). This seems to us easily solved. Weisstein's diagram for the metaccontrast interaction (1972, p. 248) is *not* symmetric, but symmetry is implicit, and it is easily obtained without making any additional as-

sumptions. Figure 5 depicts the following: Somewhere, some place in the central nervous system, two neurons receive the same excitatory input. The response of one of these neurons is faster than the response of the other, and there is cross-inhibition from these fast responders. That is, the neurons that are responding faster inhibit their successive neighbors. Because their response is fast, inhibition will be effective only if the stimulus for these fast responders is delayed. Working through the various combinations, it is immediately obvious that the network is symmetrical: All that is required for one stimulus to suppress response to another is that it follow the other. (Although the schematic shows only one line of units responding to target and one line of units responding to mask, it is assumed that there will be a collection of units responding to any stimulus.)

The Assumption of Fast Inhibition

Consider now the temporal assumption. Weisstein chose rate constants for the inhibitory factors in most of the neural units

that were faster than the rate constants for the excitatory factors. Since the inhibitory and excitatory rate constants are supposed to correspond to inhibitory postsynaptic potentials (IPSPs) and excitatory postsynaptic potentials, (EPSPs), respectively, this amounted to an assumption that IPSPs were faster than EPSPs. Although there is some evidence for fast IPSPs (Creutzfeldt & Ito, 1968), in general the assumption of IPSPs being faster than EPSPs appears tenuous.

However, there are other ways to produce fast and slow responses. With two-factor neurons, the particular rate and shape of the neural temporal response depends most critically on the relationship between the ratios of excitatory gain to excitatory rate constant versus inhibitory gain to inhibitory rate constant, and only secondarily on the rate constants themselves. One can produce a fast response in a neuron—that is (as illustrated by n_{22} in Figure 5), a response in which firing frequency reaches some criterion amplitude rapidly after stimulus onset and decreases to some criterion amplitude rapidly with stimulus offset, so that the latency of response is short, and the temporal resolution is high—by adjusting the gains in that neuron (A, B), the rate constants (a, b), or both. This can be seen by solving Equations 2 and 3 in Weisstein (1972), which would give the functions for ϵ and j , that is, for the excitatory and inhibitory factors that combine to give the neural temporal response. The solutions for a step function input, setting initial conditions to zero, are the familiar exponentials,

$$\text{Response} = \epsilon(t) - j(t) = [(A/a)(1 - e^{-at})] - [(B/b)(1 - e^{-bt})],$$

where A is the gain and a is the rate constant for the excitatory factor, and B is the gain and b is the rate constant for the inhibitory factor, and $e = 2.718$. Now, if one assumes, for instance, that $a \gg b$ (that is, EPSPs that are considerably faster than IPSPs), and $A \gg B$, such that A/a is slightly larger than B/b , a fast, sharp transient excitatory response will be produced, which will then die down to some minimal

firing slightly above spontaneous level for as long as a stimulus is maintained.

On the other hand, if one assumes $a > b$, and $A > B$, such that $A/a > B/b$, a slower, more sustained response will develop. Since a number of afferent orders are assumed to be traveled before the metacontrast interaction, one could enhance these differences by further convolution (the fast response would get somewhat slower; the slow response would get a great deal slower). Thus, there are a number of different choices of parameters that will give fast and slow responses. But one does not even have to adjust parameters relative to each other to produce Weisstein's results. One could obtain fast and slow responses simply by assuming slower conduction times in the slower responding neurons. Either (or both) of these alternate ways of producing a fast response would only minimally affect the predictions from the Weisstein network, since the narrowing of the functions and their shift towards the origin, with $T/M \rightarrow 0$, as well as the other temporal properties of metacontrast that were predicted, are due not so much to the temporal *shapes* of the slow and fast neurons (responses that are influenced by the rate constants), but to the neurons' nonlinearity (Weisstein, 1972, Equation 4, p. 246) and to the fact that the response to a stimulus, even if it were transient, would nevertheless smear out somewhat in time when the energy of the mask is high.

The nonlinearity means that the functions will first be wide and then narrow as mask energy increases; the "smearing" of the mask in time means that one will not have to delay presentation of the mask as much, as its energy increases, in order to have it interfere with the target, and thus the metacontrast functions will shift their minima toward the origin, as in Figure 1.

In short, alternate ways of producing a fast response are easily available; the focus on fast IPSPs, or "fast inhibition" (Weisstein, 1972), obscures the main point that the only thing that has to be fast in Weisstein's model is a fast excitatory response, which then inhibits a slow excitatory response. To construct this, one can choose

parameters in several ways, only one of which involves choosing faster IPSPs.

One could decide on one or another of the alternative ways of producing a fast response by drawing on the accumulating data about the properties of transient and sustained neural populations in the visual system (see below), and the equations for the two-factor neurons could be revised in light of this evidence. However, for now, since alternate constructions will yield roughly the same results, we prefer simply to note the equivalence of the different ways of looking at what was called fast inhibition.

Further Elaboration of Metaccontrast Simulation 2

Edges and blobs. The main assumption in this model is that more than one temporal weighting function is needed to characterize visual system response (i.e., fast versus slow response). So the simulation involves multiple temporal channels, two classes of which are assumed to interact. But there are spatial assumptions as well. Weisstein (1968, 1971, 1972, 1973) proposed that the units that mediated detection of a stimulus had different time constants than the units that carried information about contour or edge. She assumed, as Werner had suggested (1935) that the basis of metaccontrast was the suppression of edges or contour information, and that apparent brightness decreased as some function of the suppression of edges. Edges have long been known to play a critical role in the perception of apparent brightness (e.g., Cornsweet, 1970; Land & McCann, 1971; O'Brien, 1958; Shapley & Tolhurst, 1973).

If apparent brightness is largely dependent on edges, then "edgeless" stimuli should be more resistant to metaccontrast masking. This seems to be the case: Growney (in press) has shown that one obtains negligible amounts of metaccontrast masking when the sharp edges of a target are blurred. Masking is not as negligible when the target edges are sharp and the edges of the mask are blurred. Thus relatively "edgeless" stimuli (they are not completely edgeless, especially if one analyzes edges in terms of the amount

of high spatial-frequency components they contain) do not get masked very much.⁹

The slow response in Weisstein's model (1968, 1972) was assumed to have something to do with contour or edge. Evidence has since accumulated that high and low spatial frequencies may be processed by different temporal channels (e.g., Kulikowski & Tolhurst, 1973), and that these different temporal channels may respond with differential speed and resolution (e.g., see discussion by Breitmeyer, in press; Breitmeyer, Love, & Wepman, 1974; and Breitmeyer, Note 1; also see Hood, 1973). A high spatial-frequency response may be thought of (very roughly) as a response to edges, and a low spatial-frequency response may be thought of (just as roughly) as a response to non-edges—blobs, or large, blurred shapes, as in a defocused picture.¹⁰ Thus, given a single stimulus, the response to its edges may be slower than the response to its "blob presence." Hence, both an edge response and a non-edge response may be tied to the accumulating evidence for the presence in the visual system of channels that respond to transient, low spatial-frequency stimuli, and channels which respond to sustained, high spatial-frequency stimuli. (This direct association of fast and slow responses in metaccontrast with transient and sustained channels was first suggested by Breitmeyer

⁹ Growney (in press) has presented quite interesting evidence that the specific type of edge in metaccontrast in both target and mask can change the amount of masking obtained. While some of this is easily explained by the edge-blob scheme discussed below, some of this may require further elaboration.

¹⁰ Actually, both edges and blobs will have low frequency components, but edges will also contain more high frequency components. Edges and blobs can be specified much more precisely, of course—for any given edge or any given blob, one can calculate its spatial frequency or Fourier spectrum. Similarly, the units sensitive to edges and blobs can be defined by the spatial frequencies to which they respond. But until we know considerably more about these units' spatial frequency response—bandwidth, channel density, variation of temporal response with center frequency (see Graham, in press, for an elegant summary of our ignorance on these matters, and an excellent discussion of how to overcome it)—we lose nothing in precision by talking about edges and blobs, and we may gain in comprehensibility.

et al., 1974, and Breitmeyer, Note 1. A different approach to the relationship between masking and activity in the sustained and transient channels has been suggested by Matin, 1974, and Matin, Notes 2 and 3.)

It should be noted that transient responses do not necessarily imply speed, nor do sustained responses imply delay; however, there is evidence that the higher spatial-frequency channels may require a longer integration time than low spatial-frequency channels (e.g., see Breitmeyer, in press; Hood, 1973; Breitmeyer, Note 1), and there is neurophysiological evidence that the latencies and conduction velocities of neural populations that prefer transient, low spatial-frequency stimuli are faster than the latencies and conduction velocities of those neural populations that prefer sustained, high spatial-frequency stimuli (the Y and X cells, respectively—Enroth-Cugell & Robson, 1966. For evidence concerning conduction velocity and latency of these populations, see Cleland, Dubin, & Levick, 1971, and Fukada, 1971; for evidence that these populations may mutually inhibit each other, see Singer & Bedworth, 1973).

Breitmeyer also has gathered independent evidence from reaction times, which suggests that edges may be processed more slowly than blobs or non-edges (Breitmeyer, in press). He reported that reaction times to high spatial-frequency sinusoidal gratings are greater than reaction times to sinusoidal gratings of lower spatial frequency. These reaction-time data may reflect processing times for edges versus non-edges, since edges contain more high spatial frequencies.

Masking in General

The association of slow and fast responses with the accumulating evidence for sustained and transient, high and low spatial-frequency channels, which we will call the edge-blob scheme, has enormous integrating potential. It may make it possible to explain masking phenomena that have not yet adequately been treated; and it may help explain the relation between apparent movement and metacontrast (see also Breitmeyer et al.,

1974; Kahneman, 1967, 1968; Weisstein & Growney, 1969; Matin, Notes 2 and 3, for discussions).¹¹

As long as channels are near enough to each other to inhibit each other, the edge-blob scheme works for metacontrast, as stated: U-shaped backward masking occurs because the fast part of the response to the mask inhibits the slow part of the response to the target. This can be easily extended to nonmetacontrast designs, where U-shaped masking functions are also obtained (Purcell & Stewart, 1970; Weisstein, 1971; Turvey, 1973), that is, designs where the target and mask partially or completely superimpose on the same retinal areas. Separate channels are assumed mutually inhibiting; as they approach each other's retinal locations (including superimposition), inhibition will become stronger or remain the same, depending on one's specific spatial-frequency assumptions (e.g., Weisstein, in press). For forward masking, a U shape would be predicted if the slow part of the response to the mask inhibited the fast part of the response to the target.

Here, one might assume that since the target blob response would be inhibited (rather than the target edge response, as in backward masking), the forward analog of metacontrast—paracontrast, where the mask precedes the target—would not be as great as metacontrast, since it is the edges of a

¹¹ It should be noted that although the scheme generalizes beyond typical metacontrast backward masking, it does not (nor is it meant to) account for masking that is thought to be more peripheral—masking that decreases detection, depends primarily on overlap of the target and mask, produces monotonic functions, and so forth (see Kahneman, 1968, and Weisstein, 1972, for a review of the different types of masking). That is, although blobs might play a role in the detectability of the target, the edge-blob distinction does not commit the blobs to target detection, as might be assumed from the earlier model (Weisstein, 1968, 1972) where the basic distinction was between contour interactions and detection. Rather, masking that involves simple detection may occur as a result of different mechanisms than the ones suggested here; in any case, at the level of processing we are considering, we assume that both blobs and edges play a role in target appearance.

stimulus rather than its blob presence that seem to play such an important role in determining apparent brightness and the other features of the appearance of the target which get disrupted in metacontrast. On the other hand, if the appearance of the target is even minimally dependent on blobs as well, then, although one might not expect much paracontrast until the mask became quite strong, when the mask did become strong, forward U shapes would be expected. Finally, if the mask produced a response that was much greater than the target response, one might predict the same kind of shift of the bottom of the U that one finds in backward masking. This seems to be what happens. In Figure 1, the forward U is either absent or appears as a minor dip compared to the backward U when $T/M \approx 1$. It only reaches a respectable size as T/M decreases substantially; then, with further decreases in T/M , the bottom of the U begins to shift towards the origin. So, the edge-blob scheme might be able to account both for forward and backward Us, and for their asymmetry.

Finally, there has always been the nagging suspicion (see footnote 1) that metacontrast does not really exist. In a metacontrast design, while it is clear that target apparent brightness decreases, it is also clear that there is something going on in the location of the target, and indeed, detection measures rarely produce U-shaped metacontrast functions. Also, apparent movement is generally noticed in metacontrast, even when the target apparent brightness is negligible. The edge-blob explanation could account for both of these things: If only the edges of a target are suppressed, something indeed remains. That something would be the target blob. But since blobs produce a stronger response in the transient channels, those which prefer moving stimuli, there is reason to suppose that the continued activity of these channels during metacontrast masking would lead to the impression of movement.

CONCLUSION

A stimulus produces prolonged neural activity. The notion that different perceptual

events are associated with different portions of this activity is intriguing, and it lies at the basis of all of the theories of metacontrast discussed in this paper. The major difference in the various theories has to do with assumptions about visual system organization. Here, Bridgeman's (1971) model appears to fail in important respects, whereas Weisstein's (1968, 1972) model does not. While much of this may be due to the particulars of the way Bridgeman set up his model, some of the failure does appear intimately connected to the general notion that a single spatial and temporal channel describes visual system activity. On the other hand, Weisstein's model (with subsequent elaboration) not only offers the prospect of simulating many of the relevant spatial and temporal aspects of metacontrast, but agrees with much current investigation indicating both the presence in the visual system of multiple spatial and temporal channels (Campbell & Robson, 1968; Graham & Nachmias, 1971; Kulikowski & Tolhurst, 1973; Pantle, 1971), and also (for spatial channels at least) their mutual inhibition (Tolhurst, 1972). It is possible that by conceptualizing metacontrast in terms of multiple spatial and temporal channels, we may get a clearer idea of the variety of perceptual functions served by these different channels.

REFERENCE NOTES

1. Breitmeyer, B. G. *Implications of reaction time to sinusoidal gratings for contour forming operations in human vision*. Unpublished manuscript, 1974. (Available from the Department of Psychology, University of Houston, Houston, Texas 77004.)
2. Matin, E. *The two-transient (masking) paradigm*. Paper presented at the meeting of the Association for Research in Vision and Ophthalmology, Sarasota, Florida, May 1975.
3. Matin, E. *The two-transient (masking) paradigm*. Manuscript submitted for publication, 1975.

REFERENCES

- Alpern, M. Metacontrast. *Journal of the Optical Society of America*, 1953, 43, 648-657.
 Barlow, R. B. Inhibitory fields in the *Limulus* lateral eye. *Journal of General Physiology*, 1969, 54, 383-396.

- Breitmeyer, B. Simple reaction time as a measure of the temporal response properties of transient and sustained channels. *Vision Research*, in press.
- Breitmeyer, B., Love, R., & Wepman, B. Contour suppression during stroboscopic motion and metacontrast. *Vision Research*, 1974, 14, 1451-1456.
- Bridgeman, B. Metacontrast and lateral inhibition. *Psychological Review*, 1971, 78, 528-539.
- Brindley, G. S. *Physiology of the retina and visual pathway* (2nd ed.). Baltimore, Md.: Williams & Wilkins, 1970.
- Campbell, F. W., & Robson, J. G. Application of Fourier analysis to the visibility of gratings. *Journal of Physiology*, 1968, 197, 551-566.
- Cleland, B., Dubin, M., & Levick, W. Sustained and transient neurones in the cat's retina and lateral geniculate nucleus. *Journal of Physiology*, 1971, 217, 473-496.
- Cornsweet, T. *Visual perception*. New York: Academic Press, 1970.
- Cornsweet, T., & Teller, D. Y. Relation of increment thresholds to brightness and luminance. *Journal of the Optical Society of America*, 1965, 55, 1303-1308.
- Cox, S., & Dember, W. N. U-shaped metacontrast functions with a detection task. *Journal of Experimental Psychology*, 1972, 95, 327-333.
- Creutzfeldt, O., & Ito, M. Functional synaptic organization of primary visual cortex neurones in the cat. *Experimental Brain Research*, 1968, 6, 324-352.
- Enroth-Cugell, C., & Robson, J. G. The contrast sensitivity or retinal ganglion cells of the cat. *Journal of Physiology*, 1966, 187, 517-552.
- Eriksen, C., Becker, B. B., & Hoffman, J. E. Safari to masking land: A hunt for the elusive U. *Perception & Psychophysics*, 1970, 8, 245-250.
- Fukada, Y. Receptive field organization of cat optic nerve fibers with special reference to conduction velocity. *Vision Research*, 1971, 11, 209-226.
- Graham, N. Spatial frequency channels in human vision: Detecting edges without edge detectors. In C. S. Harris (Ed.) *Visual coding and adaptability*. Hillsdale, N.J.: Erlbaum, in press.
- Graham, N., & Nachmias, J. Detection of grating patterns containing two spatial frequencies: A comparison of single-channel and multiple-channels models. *Vision Research*, 1971, 11, 251-259.
- Growney, R. L. The function of contour in metacontrast. *Vision Research*, in press.
- Growney, R. L., & Weisstein, N. Spatial characteristics of metacontrast. *Journal of the Optical Society of America*, 1972, 62, 690-696.
- Hood, D. C. The effects of edge sharpness and exposure duration on detection threshold. *Vision Research*, 1973, 13, 759-766.
- Kahneman, D. An onset-onset law for one case of apparent motion and metacontrast. *Perception & Psychophysics*, 1967, 2, 577-584.
- Kahneman, D. Method, findings, and theory in studies of visual masking. *Psychological Bulletin*, 1968, 70, 404-425.
- Kulikowski, J. J., & Tolhurst, D. J. Psychophysical evidence for sustained and transient detectors in human vision. *Journal of Physiology*, 1973, 232, 149-162.
- Land, E. H., & McCann, J. J. Lightness and retinex theory. *Journal of the Optical Society of America*, 1971, 61, 1-11.
- Lathi, B. P. *Signals, systems, and communications*. New York: Wiley, 1965.
- Lefton, L. A. Metacontrast: A review. *Perception & Psychophysics*, 1973, 13, 161-171. (a)
- Lefton, L. A. Spatial factors in metacontrast. *Perception & Psychophysics*, 1973, 14, 497-500. (b)
- Matthews, M. L. Facilitation of letter recognition by an annulus in a metacontrast paradigm. *Quarterly Journal of Experimental Psychology*, 1974, 26, 312-323.
- Nachmias, J. Visual resolution of two-bar patterns. *Journal of the Optical Society of America*, 1968, 58, 9-12.
- O'Brien, V. Contour perception, illusion, and reality. *Journal of the Optical Society of America*, 1958, 48, 112-119.
- Pantle, A. Flicker adaptation, I: Effect of visual sensitivity to temporal fluctuations of light intensity. *Vision Research*, 1971, 11, 943-952.
- Purcell, D., & Stewart, A. U-shaped backward masking functions with non-metacontrast paradigms. *Psychonomic Science*, 1970, 21, 361-363.
- Ratcliff, F. *Mach bands: Quantitative studies on neural networks in the retina*. San Francisco: Holden-Day, 1965.
- Schiller, P. H., & Smith, M. Detection in metacontrast. *Journal of Experimental Psychology*, 1966, 71, 32-46.
- Sekuler, R. W. Review of *Handbook of sensory physiology*, Vol. 7/4, *Visual psychophysics* by D. Jameson & L. Hurvich (Eds.). *The American Journal of Psychology*, 1973, 86, 876-886.
- Shapley, R. M., & Tolhurst, D. J. Edge detectors in human vision. *Journal of Physiology*, 1973, 229, 165-183.
- Singer, W., & Bedworth, N. Inhibitory interaction between X and Y units in the cat lateral geniculate nucleus. *Brain Research*, 1973, 49, 291-307.
- Stigler, R. Chronophotische studien uber den umgebungskontrast. *Pflügers Archiv für die Gesamte Physiologie*, 1910, 134, 365-435.
- Sturr, J. F., & Battersby, W. S. Neural limitations of visual excitability. VIII. Binocular convergence in cat geniculate and cortex. *Vision Research*, 1966, 6, 401-418.
- Tolhurst, D. J. Adaptation to square-wave gratings: Inhibition between spatial frequency channels in human visual system. *Journal of Physiology*, 1972, 226, 231-248.

- Turvey, M. T. On peripheral and central processes in vision: Inferences from an information-processing analysis of masking with patterned stimuli. *Psychological Review*, 1973, 80, 1-52.
- Weisstein, N. A Rashevsky-Landahl neural net: Simulation of metacontrast. *Psychological Review*, 1968, 75, 494-521.
- Weisstein, N. W-shaped and U-shaped functions obtained for monoptic and dichoptic disc-disc masking. *Perception & Psychophysics*, 1971, 9, 275-278.
- Weisstein, N. Metacontrast. In D. Jameson & L. Hurvich (Eds.), *Handbook of sensory physiology* (Vol. 7, part 4: *Visual psychophysics*). Berlin: Springer-Verlag, 1972.
- Weisstein, N. Beyond the yellow Volkswagen detector and the grandmother cell: A general strategy for the exploration of operations in human pattern recognition. In R. L. Solso (Ed.), *Contemporary issues in cognitive psychology: The Loyola Symposium*. Washington, D.C.: V. H. Winston, 1973.
- Weisstein, N. Masking and unmasking of distributed representations in the visual system. In C. S. Harris (Ed.), *Visual coding and adaptability*. Hillsdale, N. J.: Erlbaum, in press.
- Weisstein, N., & Growney, R. L. Apparent movement and metacontrast: A note on Kahneman's formulation. *Perception & Psychophysics*, 1969, 5, 321-328.
- Weisstein, N., Jurkens, T., & Onderisin, T. Effect of forced-choice vs. magnitude-estimation measures on the waveform of metacontrast functions. *Journal of the Optical Society of America*, 1970, 60, 978-980.
- Werner, H. Studies on contour: 1. Qualitative analyses. *The American Journal of Psychology*, 1935, 47, 40-64.

(Received December 16, 1974)

Neural Stability, Sparing, and Behavioral Recovery Following Brain Damage

T. E. LeVere
North Carolina State University at Raleigh

The present article discusses the possibility that behavioral recovery following brain damage is not dependent on the functional reorganization of neural tissue but is rather the result of the continued normal operation of spared neural mechanisms. In support of this position, data are reviewed which indicate that the nervous system, like other biological systems, is quite stable and incapable of the plasticity required to mediate functional reorganization. Within this frame of neural stability, representative research concerned with recovery of function is reviewed, and the suggestion is offered that the phenomenon of behavioral restitution following brain damage is not only compatible with neural stability and sparing, it is in direct support of it. Finally, several possible mechanisms capable of mediating behavioral recovery under the aegis of neural stability and sparing are discussed.

One of the more widely accepted propositions describing neural functioning is summarized by the following statement: "The initial effects of neurological insult are more severe than the subsequent long-term effects." The statement describes the belief that some behavioral restitution following brain damage is the rule and not the exception. The breadth of agreement which the statement commands reflects the behavioral and anatomical range over which it appears to be valid. Behaviorally, this range extends from more basic innate activities, such as sleeping (Adametz, 1959; Lourie, Vanasupa, & O'Leary, 1960) and feeding (Teitelbaum & Epstein, 1962) to learned behaviors based on relatively simple visual (Lashley, 1935), auditory (Raab & Ades, 1946; Rosenweig, 1946), and tactile cues (Finger, Marshak, Cohen, Scheff, Trace, & Niemand, 1971) to ever increasingly complex behaviors involving patterned visual (Mize, Wetzel, & Thompson, 1971; Spear & Braun, 1969; Winans, 1967) and auditory stimuli

(Butler, Diamond, & Neff, 1957; Meyer & Woolsey, 1952). From an anatomical view, the range of neural centers that apparently exhibit some degree of functional restitution is also quite extensive and extends from the level of the spinal cord and brain stem to diencephalic centers and neocortex, as exemplified by the research noted above.

On the basis of these and similar data, the present discussion accepts the validity of recovery of function following brain damage and the influence of certain parameters, such as infancy and serial damage, and from this point will consider what process or processes may underwrite the phenomenon itself. Neither the histological accreditation of the phenomenon nor the delineation of the various operational parameters which influence the process will be considered. These issues have been addressed by Rosner (1970, 1974) and little can be added to these elegant reviews.

At the outset, it is perhaps best to state explicitly that the primary interest of the present discussion is the recovery of specific behavior patterns that may be disrupted by damaging specific brain areas. A classic example, in the tradition of K. S. Lashley, is the recovery of a two-choice brightness discrimination following striate decortication. The general paradigm involves training subjects, usually rats or cats, to select

The author's research was supported by one or more of the following: National Institutes of Health, Biomedical Support Grant RR07071; National Science Foundation, Institute Grant for Science GU-3610.

Requests for reprints should be sent to T. E. LeVere, Department of Psychology, North Carolina State University, Raleigh, North Carolina 27607.

one of two brightnesses in a two-choice discrimination box. After the animal has mastered the discrimination, it is subjected to removal of some portion of its brain—in the present instance the striate neocortex—and the behavioral effects of this neural damage are assessed by later testing in the discrimination box. Postoperatively, this testing typically indicates that the striate animal is initially unable to perform the brightness discrimination but with some practice is able to reestablish the preoperative discriminatory behavior.

Traditionally, these results have been interpreted, more often than not, as indicating the occurrence of functional reorganization. With regard to the above example, this concept suggests that initially the brightness behavior is mediated by neurological mechanisms within the striate neocortex. When the striate neocortex is destroyed, these neurological mechanisms are lost, and the behavior they supported is also necessarily lost. The animal is able to recover the behavior, however, because on an operational level, the nervous system's response to the neurological insult is a functional reorganization of its surviving component parts so that other neural centers may come to function in place of the destroyed neural tissue. The postoperative training thus allows reestablishment of the brightness discrimination, because these functionally reorganized neural centers have taken over the operational properties of the lesioned brain area. In a general sense, this sort of explanation, while somewhat gross, has a nontrivial elegance in its simplicity and gains a certain amount of stature from the plasticity the nervous system exhibits with respect to experiential modification.

However, there are two serious (fatal, from the present point of view) problems with the notion of functional reorganization as it is evoked to account for behavioral recovery following brain damage. First, the concept suggests a characteristic, "design plasticity," which is unique among the properties of biological systems. Design plasticity is here used to denote a functional change in the operational mode of a particular neural center or area. This is not, of course,

to deny experiential modification or learning, since this sort of plasticity is well within the design characteristics of neural tissue and moreover is not a property unique to the nervous system. Rather, the problem rests on the plasticity necessary to underwrite functional reorganization, which is something much more profound and quite different, requiring the modification of the design characteristics of the system itself—a property which, if it were to exist, would be most unique. Second, and more important, this design plasticity, which is a requisite to functional reorganization, is incredibly difficult to support with any sort of direct empirical evidence. That is, while the neuronal plasticity suggested to account for learning is often precipitated by the direct observation of some particular physiological response of neural tissue, the concept of functional reorganization does not enjoy any such complementary justification. In point of fact, the only evidence for functional reorganization is recovery of function, and this is obviously quite circular.

In contradistinction to the notion of functional reorganization, the present paper argues that the nervous system is not as unique as this concept suggests. Rather, it is suggested that the mature nervous system, like other mature biological systems, is quite stable after some critical point in its embryological and/or neonatal development. Accepting this, the present paper then proposes that recovery of function following brain damage depends on the sparing of the neural mechanisms responsible for the behavior patterns prior to the neural damage. In other words, recovery of function following brain damage does not reflect the functional reorganization of neural tissue but rather the survival of the neural tissue that mediated the behavior patterns prior to the neurological lesion.

The remainder of this paper attempts to support these notions of functional stability and neurological sparing as determinants of recovery of function following brain damage. To do this, two things must be accomplished. First, it must be shown that the mature nervous system is functionally stable. Second, a case must be made that

behavioral recovery represents no more or no less than the continued normal operation of the surviving brain centers. In this latter instance, it is specifically suggested that the recovery of behaviors disrupted by neural lesions is the result of the fact that the neural mechanisms supporting these behaviors were never really destroyed in the first place.

STABILITY WITHIN THE MATURE NERVOUS SYSTEM

The evidence suggesting that the mature nervous system is functionally quite stable stems principally from three sources: (a) neural regeneration, (b) early experience, and (c) compensatory neural sprouting. As will be apparent, all of this research describes very dynamic aspects of the nervous system and yet, at the same time, provides quite convincing evidence that these dynamic aspects do not result in the functional reorganization of established neural centers. And, as should be apparent, this failure to support functional reorganization is in precisely those areas of research that should demonstrate it if it were a viable process.

Neural Regeneration

While regeneration following neural damage is not ordinarily considered a property of the mature mammalian central nervous system, it is readily observed in the mammalian peripheral nervous system and in the central nervous system of lower vertebrates. Of critical import for the present discussion is not that there is regeneration but rather that regeneration, when it occurs, is quite specific. This was recognized early (Dustin, 1910; Held, 1905; His, 1889—all cited in Grossman, 1967) and generated considerable controversy concerning the mechanisms involved. Mechanical theories relying on stereotropic factors (Haidenhain, 1911—cited in Grossman, 1967) were advanced but proved somewhat inadequate in the face of data indicating that the outgrowing amoeboid nerve endings did not always follow the path of least resistance (Grossman, 1898—cited in Grossman, 1967). Neurotropic theories involving

chemical specificity, as initially proposed by Ramón y Cajal (1928/1959) and later supported and refined by the classic research of Weiss (1936) and Sperry and his collaborators (Sperry, 1947, 1950, 1955, 1958; Sperry & Hibbard, 1968), to name but the more familiar, now appear considerably more likely. However, whatever the mechanisms, and the options have been carefully considered by other authors, the point remains that whether it is mechanical, chemical, or some thing such as neurobio-taxis (Kappers, 1917), neural regeneration is not a haphazard affair. On the contrary, it is quite specific and when it occurs, as when eyes are rotated, an incredible degree of functional stability is maintained, even when this stability is contrary to adaptive behavior (also see recent research of Mark and his associates [Cass, Sutton, & Mark, 1973; Mark & Marotte, 1972; Mark, Marotte & Mart, 1972; Marotte & Mark, 1970a, 1970b]). If the mature nervous system is dynamic or plastic enough to allow functional reorganization following brain damage, it is difficult to understand this degree of regenerative specificity.

Early Experience

In contrast to the specificity observed during regeneration in the mature nervous system, the nervous system of the neonate appears to be capable of considerably plasticity. The extent of this plasticity is such that rather large changes in the functional organization of the nervous system as well as more concrete anatomical changes, for example, the thickness of the cortex, dendritic branching, and the like, can occur in response to a variety of experimental manipulations. This is, of course, just the sort of neural responsiveness postulated by the concept of functional reorganization to account for behavioral recovery following brain damage. However, contrary to the notion of functional reorganization, once the individual passes some critical period and approaches maturity, this plasticity is lost and the nervous system becomes very stable. Moreover, the data indicate that this stability is maintained in spite of extensive behavioral experience and procedures specifically de-

signed to induce neural reorganization and restitution of function.

Early research in this area comes from the eloquent studies of Wiesel and Hubel and their associates (Dews & Wiesel, 1970; Hubel & Wiesel, 1965, 1970; Wiesel & Hubel, 1963). The general procedure used by these investigators was to restrict visual experience for a short period of time during the first few months of life by suturing closed the lids of one eye. The effects of this visual deprivation were then assessed by analyzing single-unit responses in the striate neocortex when the animal had developed to maturity. In the normal cat (that is, the cat reared without visual deprivation) approximately 80% of the single units were binocularly responsive. In the deprived cat, however (that is, the cat reared for a short time with one of its eyes sutured shut), this binocular responsiveness dropped to approximately 7% or less. With regard to the stability of the nervous system, it is important to note that once past the 7th to 10th postnatal week, these changes are quite permanent and show little modification, even when the cat is provided with up to 4 years of normal binocular visual experience. It would seem, if the nervous system were capable of functional reorganization, that this amount of normal visual experience should allow considerable recovery. That it does not would appear to argue that neural centers, once mature, are functionally stable. This stability is further attested to by the fact that similar visual deprivations in adulthood, even of 1 year's duration, are without noticeable effect on cortical electrophysiology if the animal enjoyed normal binocular visual experience during its development.

Other investigators have reported results similar to those of Hubel and Wiesel, results that are noteworthy because they did not involve totally depriving the individual of visual stimulation (Blakemore & Cooper, 1970; Hirsch & Spinelli, 1971; Pettigrew & Freeman, 1973). For example, Blakemore and Cooper reared cats in darkness during the early developmental period with the exception of daily 5-hour exposures to either horizontal or vertical stripes. When

the kittens were 5 months old, they were allowed normal visual experience and underwent both behavioral and electrophysiological testing. After some minimal adaptation period, the subjects appeared to be visually quite competent with the exception of being virtually blind to stimuli oriented 90° to their early stimulation. Single-unit recordings from the striate neocortex of these animals detected the normal amount of binocularly responsive cells, except that units responding to lines oriented at right angles to the kittens' early visual experience were conspicuously absent. And it is critically important to note that this recording was accomplished after the subjects had undergone some amount of behavioral testing, which should have provided time and specific stimulation for visual reorganization if it were to occur. As Blakemore and Cooper (1970) stated, "Evidently the visual experience of these animals in early life has modified their brains, and there are profound perceptual consequences" (p. 478). To this, the present argument would only note that these changes appear permanent.

Similar results have been obtained when the visual experience has been restricted to small point sources of light (Pettigrew & Freeman, 1973), as well as when different types of restrictions have been placed on each individual eye (Hirsch & Spinelli, 1971). Recently, Freeman and Thibos (1973) have extended these findings to humans by correlating visually evoked responses to meridional resolution in astigmatic individuals. While the mechanisms underwriting this developmental plasticity and adult stability are presently somewhat speculative and extend from postulations of cortical units dropping out due to lack of stimulation to the tuning of units to correspond to the environment in which the individual is reared, the fact remains that whatever occurs, when it occurs, it is permanent. As implied by all of the previously discussed research and specifically tested by Dews and Wiesel (1970) and Ganz and Haffner (1974), normal visual experience does not rectify the conditions established during the sensitive developmental period. Even suturing the eyelid of the originally

undeprived eye, thus compelling the animal to use the previously deprived eye, has only minimal effects on overt visual behavior and is virtually without influence on cortical electrophysiology. What one is then forced to accept on the basis of these data is that subsequent to maturation, the nervous system is functionally stable. Given this sort of neural stability, it would seem somewhat inappropriate to suggest that behavioral recovery following brain damage is the result of neural centers changing or reorganizing their functional mode of operation.

Neural Sprouting

Further evidence that the mature mammalian nervous system is functionally stable, but not static, is provided by the research concerned with neuromorphological plasticity. Similarly to the early-experience research, the data present a picture of the nervous system as a biological system capable of growth and change, but only within a relatively restricted frame of reference—a frame which, in the mature individual, does not encompass functional reorganization. It must be emphasized, however, that what is at issue is not whether the peripheral nervous tissue may reinnervate sensory or motor mechanisms and thus precipitate behavioral recovery, nor is it denied that reinnervating may even occur centrally to provide some measure of behavioral recovery. In both of these cases, the effect is not one of reorganization but rather one of reinstitution, and this is well within the assumption of neural stability espoused by the present report.

From the functional point of view, the data concerned with neural sprouting indicate, as in the early-experience data, a highly plastic nervous system during early neonatal development but a nervous system that rapidly stabilizes during postnatal development. I will first consider the research concerned with neural sprouting in the mature animal.

Neural Sprouting: The Mature Nervous System

With regard to the mature nervous system, the major question is whether the

manipulations that induce regenerative or collateral neural sprouting, as the terms are defined by Stenevi, Björklund, and Moore (1972), will produce generalized or restricted neural sprouting. If the observed neural sprouting following neurological insult is generalized (i.e., it includes neural centers or fiber processes normally not part of the system affected by the lesion) then a primary condition for functional reorganization is met and the notion may gain considerable support. If, however, the neural sprouting is restricted to the areas directly related to the neural insult, functional reorganization is considerably less likely. In this latter case, as previously suggested, the effect of the neural sprouting would be better characterized as functional reinstitution, not functional reorganization, and as such would be compatible with the notion of neural stability.

By and large, the data suggest that compensatory neural sprouting is quite restricted. The recent research of Goodman, Bogdasarian, and Horel (1973) may be taken as a specific example. This research was generated as an extension of earlier experiments (Goodman & Horel, 1966) indicating that compensatory neural sprouting in response to ablation of the striate neocortex occurred in only 2 of the possible 13 areas suffering deafferentation subsequent to the neocortical insult. At the risk of oversimplifying an eloquent conceptualization, the methodology of Goodman et al. (1973) may be summarized as one that tested for generalized compensatory sprouting by optimizing the conditions necessary for its occurrence. This was accomplished by partially deafferenting the central visual areas of 3-month-old rats by unilateral enucleation. Some 16 months later, the remaining eye was removed and after 5 days the tissue prepared for analysis of degenerating fiber processes, using the Laidlaw modification of the Nauta-Gygax method. The arguments were that the early partial deafferentation should produce a situation where competition from systems other than the visual system would be minimized because of the remaining retinofugal fibers. This should, in turn, allow gen-

eralized compensatory sprouting in all deafferented sites. The sprouting that was observed, however, was only at the sites of normal termination of the retinofugal fiber processes remaining after the initial enucleation. Increased distribution, which would seem to be a minimal but critical requirement for functional reorganization, was not observed at any of the visual centers, with the possible exception of the superior colliculus. Thus, if the physiological concomitants of functional reorganization occurred, they appeared quite restricted. To suggest that this limited increase in distribution might be responsible for behavioral recovery under the aegis of functional reorganization is, in most generous terms, stretching an empirical finding.

In candor, however, it must be noted that some early research concerned with compensatory sprouting has indicated some amount of generalized innervation (McCouch, Austin, Liu, & Liu, 1958; Ramón y Cajal, 1928/1959; Rose, Malis, Kruger, & Baker, 1960), although more recent research has suggested that the neural sprouting is very restricted (Cunningham, 1972; Goodman et al., 1973; Goodman & Horel, 1966; Lynch, Matthews, Mosko, Parks, & Cotman, 1972; Moore, Björklund, & Stenevi, 1971; Raisman, 1969; Schneider, 1970; Stenevi et al., 1972; Wall & Egger, 1971). Moreover, and most important, anomalous connections resulting from the compensatory neural sprouting have *never* been reported in experiments using mature subjects (Cunningham, 1972; Cunningham & Lund, 1971; Goodman & Horel, 1966; Liu & Chambers, 1958; Lynch et al., 1972; McCouch et al., 1958; Moore et al., 1971; Raisman, 1969; Wall & Egger, 1971). These latter results are perhaps worth emphasizing, because it would seem that anomalous compensatory sprouting would be very strong neural anatomical evidence for functional reorganization, should it occur in the mature nervous system. That it does not would seem to suggest that behavioral recovery must hinge on processes not involving the vicarious functioning of neural centers not normally involved in the behavior in question.

Neural Sprouting: The Neonate

In distinction to manipulations performed on the mature nervous system, compensatory sprouting following neonatal lesions produced quite a different pattern of results. Here, as with the effects of early experience, the picture is one of general neural plasticity.

For example, the normal rat's visual system is almost totally crossed with only a limited number of ipsilateral retinofugal fiber processes, and these have a very restricted central distribution (Hayhow, Sefton, & Webb, 1962). If, however, unilateral enucleation is performed during the first 10 days of postnatal life, there is a significant increase in the number of unilateral retinofugal projections (Lund, Cunningham, & Lund, 1973). Moreover, these ipsilateral fiber processes originate from retinal areas normally producing only contralateral projections and terminate in geniculate and collicular areas normally receiving only processes from the contralateral eye. Thus, it would appear that, similar to the early-experience research, there is considerably greater flexibility in the developing nervous system than there is in the mature nervous system.

Perhaps a more dramatic demonstration of this sort of developmental plasticity is provided by Schneider (Schneider, 1970, 1973), who was concerned with neonatal brain stem lesions. This research not only confirmed the existence of anomalous synaptic connections but also provided convincing evidence that these synaptic connections were operational. At first glance, the existence of operational anomalous synaptic connections would seem favorable to the concept that behavioral recovery is mediated by functional reorganization, but, as will be more fully discussed below, an analysis of the resulting overt behavior indicates that just the opposite is true. The general experimental procedure used by Schneider was to ablate, at birth, the superficial layers of one superior colliculus in a number of golden hamsters. The animals were then allowed to grow to maturity with normal visual stimulation and then were given a series of behavioral tests to determine the effects of

the early superior colliculus lesions. Following this, the animals underwent the usual procedures to determine the extent of compensatory neural sprouting, which involved unilateral enucleation followed, 5 days later, by histological analysis using various modifications of the Nauta method to assess degenerating fiber processes.

This analysis indicated rather striking abnormalities in the distribution of retinofugal fiber processes and terminations. Principally, there was evidence for an abnormally high number of fiber terminations in the ventral lateral geniculate nucleus. Additionally, and somewhat more interestingly, there was evidence of fiber termination in the posterior lateral nucleus of the thalamus as well as the lower, undamaged layers of the lesioned superior colliculus, both of which do not normally receive such fibers. Moreover, not only did the retinofugal fiber processes reach the lower, undamaged layers of the lesioned superior colliculus, but some also decussated and invaded the opposite, unlesioned superior colliculus, where they competed for terminal space with the fiber processes normally afferent to that area. Thus, the impression that the developing nervous system is highly plastic and matures in accord with the available sensory input and, in the present instance, available neural mechanisms is again reinforced.

But this plasticity of the developing nervous system, as in some of the regeneration data, falls short of being functionally adaptive. For example, behavioral testing of those subjects subsequently shown to possess the abnormal decussating fibers to the unlesioned superior colliculus showed rather strange visual orienting behavior. In these animals, visual stimuli presented to the eye *abnormally* projecting to the unlesioned superior colliculus often led to orienting responses *away* from the visual stimulus. That is, the animal would respond as though the opposite eye were receiving the visual stimulation. While this clearly suggests that the anomalous synaptic connections were operational, it also, at the same time, clearly suggests that the anatomical plasticity of the retinofugal fibers and the neural reorganization observed by Schneider (1970, 1973)

did not yield correlative behavioral restoration. I would suggest that the reason that this behavioral restoration did not occur is because the plasticity of the retinofugal pathways was not associated with a similar plasticity within the mesencephalic termination area. Rather, at birth the superior colliculus and perhaps other brainstem visual centers are probably close to maturity, and, like other areas reaching maturity after birth, they are functionally stable. The fact that the anomalous projection did not cause any sort of functional reorganization of the mature surviving superior colliculus is critical and clearly indicates a lack of functional reorganization, even when the conditions seem optimal for its occurrence (i.e., operational anomalous input). More simply put, the overt behavior patterns exhibited by lesioned animals appear to be a function of the neural areas spared by the neurological insult, and once mature, these neural areas do not reorganize in any functional sense but simply continue to operate in accord with the principles established at maturation.

In summary, what the research concerned with neural sprouting, along with the regeneration data and early-experience data, suggests is that the nervous system, like other biological systems, is plastic and malleable during its developmental stages. However, once development has run its course, the nervous system, again like other biological systems, becomes quite stable. And in functional terms, this stability is maintained in the face of rather severe physiological manipulations. What the remainder of the present report attempts to show is that this degree of stability is not incompatible with the observation of behavioral recovery following neurological insult. In fact, much of the data concerned with such behavioral recovery is not only compatible with the notion that neural centers are functionally stable, but provides rather striking evidence to support it.

SPARING AS A CONCEPT OF BEHAVIORAL RECOVERY FOLLOWING BRAIN DAMAGE

Accepting, if only for the moment, a stable nervous system incapable of functional reorganization, how is it possible to explain

behavioral recovery following neurological insult? The possibility I would like to suggest is based on sparing. To reiterate, the thrust of the present paper is that the nervous system is not unique among the various biological systems supporting life, and, like other biological systems, operational recovery following damage is dependent on what survives and what continues to function in its normal manner. From this point of view, behavioral recovery following brain damage is simply the result of the neural mechanisms supporting the recovered behaviors never being destroyed.

In support of this position, one might consider first the sensorimotor recovery that may be observed following lesions that sever peripheral neural fibers. Initially, the ability to sense and to move the affected limb is lost, and by common tests and definitions, the deficit is, in fact, neurological. However, after some period of time, calculated at about 3 mm per day from the site of the lesion to the location of the affected area, sensory and motor functions may often recover. Clearly, the case is one of functional recovery involving neural tissue. But just as clearly, the behavioral recovery does not involve any sort of functional reorganization, where one neural center changes its mode of operation in response to the destruction of some other neural center. The process is simply one of regeneration and reinnervation, and little else. Obviously, as previously pointed out, considerable work remains to be done to detail precisely the mechanisms which govern the reinnervation. However, the fact nonetheless remains that the sensorimotor behavior was recovered because, in the real sense, the neural mechanisms underlying it survived the neurological insult.

If one turns to behavioral recovery involving the central nervous system, the question becomes whether or not a similar proposition may be valid. From the arguments of the present report, the answer is yes, and it follows from a series of studies suggesting that prelesion behavior patterns often survive neurological insult even though they may not be initially observed in the subject's overt behavior immediately after the brain lesion. And it is of some importance to note that the

range of behaviors over which this appears to be true is as impressive as the range of behaviors which may demonstrate recovery subsequent to neurological damage.

One might consider first the reflexive behavior of visual placing. The usual procedure to test visual placing is to hold an animal, for example a cat, either above or below a table and slightly away from the edge. The animal is moved slowly up or down toward the edge of the table, with the result that when the edge comes into the animal's field of vision, the cat promptly and accurately places its forepaws on the table. Early data from Bard and Brooks (1934) indicated that after removal of the animal's visual neocortex, visual placing was lost and did not appear to recover even with extensive postoperative testing. These results were later replicated by Meyer, Horel, and Meyer (1963) and were generally attributed to the animal's inability to detect edges (i.e., a dysfunction of pattern vision). However, Meyer et al. (1963) and more recently Meyer, Meyer, and Ritchie (cited in Meyer, 1974) noted that if the cat was tested under the influence of the drug dl-amphetamine, visual placing was reinstated and remained as a viable behavior pattern so long as the drug state was maintained. Obviously, these data provide rather salient evidence that the placing behavior disrupted by visual neodecortication was not destroyed, even though it was not easily detectable in the subject's postoperative behavior. Moreover, and more important, it is difficult to attribute the return of visual placing under the influence of dl-amphetamine to any sort of neural reorganization. In the first place, behavioral recovery is perfectly correlated with the assimilation and systemic action of the drug, and in the second place, it is difficult to understand how the neural stimulant would cause a functional reorganization that is not permanent.

Another reflexive behavior disrupted by neurological insult but apparently not destroyed, lost, and recovered through the functional reorganization of neural centers is visual orienting. The phenomenon that demonstrates this was described by Sprague (1966) and involves a unilateral lesion of

the visual neocortex followed sometime later by destruction of the contralateral superior colliculus. Subsequent to unilateral extensive visual neodecortication (occipitotemporal cortex from splenial sulcus to rhinal fissure, or essentially the posterior half of the available cortex), cats are contralaterally hemianopic with regard to a number of visual behaviors, such as blinking in response to objects thrust into their field of vision, following moving objects, avoidance of obstacles, visual placing, and depth perception. However, the removal of the superior colliculus on the opposite side of the brain immediately restores visual functions to the hemianopic field. Here, as with the research of the Meyer group, the recovery of the visual behavior is difficult to align with the concept of functional reorganization of neural centers, for similar reasons. Paramount among these reasons is the fact that the recovered behavior appears immediately following the subcortical lesion. Furthermore, it is difficult to understand how a subcortical lesion might spontaneously reorganize neural centers to the extent that areas previously unable to support the visual behavior are now able to do so. Following the arguments of the present report, a more plausible explanation of these results would be that the visual behaviors were recovered because the neural mechanisms that control them survived the initial visual lesion, and the second lesion simply enabled their expression. This, as is discussed below, is quite similar to the interpretation offered by Sprague (1966) in his analysis of the data.

Turning to more complex behaviors outside the visual realm, one may consider the recovery of feeding behavior following hypothalamic lesions. The lateral hypothalamic syndrome, as it is now called, summarizes the aphagia and adipsia which typically result from bilateral destruction of the lateral hypothalamic nuclei. While the aphagia and adipsia are quite severe initially, the lesioned animals may regain nearly normal feeding behavior if provided with rather extensive postoperative care. Teitelbaum and Epstein (1962) have detailed with great precision the progress of this re-

covery, as the animal successfully progresses from tube feeding to accepting highly palatable foods to eventually eating dried rat pellets and drinking on its own. Considering the time course and the required postoperative therapy, this is apparently a clear example of functional reorganization, where neural centers other than the lateral hypothalamus are operationally reprogramed to mediate the feeding behavior. The observation that postnatal development of neonatal feeding may parallel this recovery process (Teitelbaum, Cheng, & Rozin, 1969) is further support for this contention. The evidence would, in fact, seem compelling, except for one peculiar problem reported by Powley and Keeseey (1970). These researchers noted that if the rat was starved to 80% of its body weight *prior* to the lateral hypothalamic ablation, there was no postoperative period of aphagia and adipsia. Moreover, if the animals were starved beyond 80% body weight, the postlesion behavior was one of hyperphagia. While Powley and Keeseey attribute these results to a change in the animal's weight "set point," what is important for the position of the present paper is that it is difficult to align these results with the notions that the recovery of feeding behavior is a result of the functional reorganization of neural tissue—unless one is willing to suggest that starvation can precipitate functional reorganization. Again, behavioral recovery would appear to occur simply because the neural mechanisms controlling the behavior survived the neural lesion.

The two-choice brightness discrimination, as described by Lashley (1935), may be used as the general paradigm to inquire if the arguments presented so far may generalize to the recovery of learned behaviors. The general procedure requires that an animal, typically a rat, be trained in a two-choice discrimination box to avoid and/or escape footshock. The discriminative cues differ in brightness, with the darker cue usually black and the lighter cue usually white. The cues are generated either by reflecting light off painted panels (Thompson-Bryan box) or by back lighting translucent panels (Yerkes box). Behavioral

training consists of a number of daily trials in the apparatus where the animal is allowed to avoid/escape shock by responding to the cue leading to a safe goal area. After the subject has mastered the discrimination to some preset criterion, say 9 correct choices out of 10 trials, it is subjected to bilateral posterior neodecortication to include those areas efferent to the dorsal lateral geniculate nucleus, as described by Lashley (1934) and more recently mapped by Adams and Forrester (1968) and Montero (1973). After a suitable recovery period, on the order of 2 weeks to a month, the animal is returned to the discrimination box and retested on the preoperatively mastered brightness discrimination.

The typical result of this retesting is that the animals initially perform at chance and require approximately the same number of trials to reestablish the brightness discrimination as were required to master it during the original preoperative training. Thus, the procedure illustrates that the learned two-choice brightness discrimination may be disrupted by striate neodecortication but also that it is not beyond the reach of the operated animal. The question of interest is whether the recovery of the brightness habit indicates some sort of neural reorganization or, alternatively, the sparing and reactivation of the neural mechanisms that controlled the behavior pattern preoperatively.

The first indication that the neural mechanisms underlying discrimination behavior may not be lost following a striate lesion comes from the research of Braun, Meyer, and Meyer (1966). These investigators studied the possibility that the drug amphetamine, which facilitates visual placing in the striate cat, might also facilitate the restoration of two-choice discriminatory behavior in rats. Their procedure was elegantly simple and followed the classic Lashley (1935) paradigm, but with the important exception that some of the rats were postoperatively retrained under the influence of the drug amphetamine, while other groups received injections of saline or were simply retrained. The results indicated that the drug amphetamine facilitated the reacquisi-

tion of the brightness discrimination, while saline injections had virtually no effect relative to the control animals who simply relearned the brightness discrimination.

By itself, this result is somewhat unimpressive, since it is possible that the drug amphetamine might facilitate acquisition independent of any other condition of the experiment. However, a subsequent experiment reported by Braun et al. (1966) suggested that the facilitative effects of amphetamine were dependent on whether or not the lesioned animal had learned the brightness discrimination preoperatively. In other words, the drug amphetamine was facilitative only because there was a previously established behavior pattern to facilitate. Since the drug was without effect if the animal had not had preoperative training, it would seem obvious that something of the preoperative brightness habit must survive striate neodecortication. And should one argue that amphetamine really facilitates functional reorganization—not reactivation—there is the nagging question of why this facilitation of functional reorganization should occur only when the subject was preoperatively trained on the brightness discrimination. Again, it would appear that the postoperative capabilities of a subject are dependent on what is spared by the neural insult and not what is reorganized after the neural insult.

However, as is often noted by the skeptical, there are two effects of drugs on behavior, "those you are aware of and those you are not aware of." To argue the case for sparing as the process of behavioral recovery, it might be more convincing to restrict the experimental techniques to those involving only behavioral procedures. This was accomplished recently by LeVere and Morlock (1973). These researchers trained rats on a brightness discrimination task and subsequent to mastery subjected the animals to extensive ablations of the posterior neocortex. Subsequent to recovery, the lesioned animals were divided into groups, which were either retrained on the original brightness discrimination or retrained on the reversal of the original brightness discrimination. The argument was that if the neural traces of the preoperatively established be-

havior survived the posterior neocortical ablation, then retraining on the reversal of the original brightness discrimination should prove quite difficult. This is so because the spared neural mechanisms would be the exact antithesis of what is required postoperatively of the reversal group.

The results indicated that the postoperative reversal group was indeed significantly impaired, and since this impairment could not be attributed to either differential brightness preferences (if such exist) or differences in neural damage, LeVere and Morlock concluded that not only were the preoperative neural mechanisms spared, but they were also apparently influential in the postoperative recovery of the visually guided behavior. Subsequent research by these same investigators (LeVere & Morlock, 1974) has indicated that the result is not restricted to the simultaneous brightness discrimination paradigm but also occurs when rats are trained on a successive brightness discrimination task. And again, the data are incompatible with the concept that behavioral recovery is mediated by functional reorganization. The postoperative behavior patterns available to the individual are simply those behavior patterns that are dependent upon neural mechanisms spared by the neurological lesion.

Thus it would appear that the evidence suggests that behavioral recovery following neurological insult may often be attributed to the possibility that the underlying neural mechanisms were not lost in the first place. The range over which this conclusion appears to hold true would suggest that it is quite general and not restricted to reflexes, basic homeostatic behaviors, or learned behaviors. To maintain, on the other hand, that behavioral recovery is the result of some unspecified, unique characteristic of neural tissue allowing for functional reorganization is not just an unnecessary complication but rather, and more seriously, a denial of the experimental data. Put another way, it is difficult to appreciate the need for the functional reorganization of neural centers to enable the vicarious *duplication* of neural mechanisms apparently not eliminated by neurological destruction.

A Word on Learning

It must be pointed out that the present arguments do not deny or disallow functional recovery through a process of learning. In fact, the possibility that learning may precipitate functional recovery following neurological insult is directly compatible with the notions of functional stability and sparing, as these concepts are developed in the present paper. In this instance, what is spared is not a specific neural mechanism and its associated behavior pattern but rather a process or capacity. The only rub is that it is difficult, in some definitional sense, to consider learning or relearning as a mechanism for the recovery of specific behavior patterns. Learning from the point of view of the individual is simply learning, whether it occurs prior to or subsequent to brain damage. Moreover, there is the distinct possibility that the behavior "recovered" postoperatively might be quite different in detail from the preoperative behavior, because, for example, the individual may be relearning on the basis of different environmental cues (Bauer & Cooper, 1964; Bland & Cooper, 1970; Cooper, Blochert, Gillespie, & Miller, 1972; Goodale & Cooper, 1965). In this case, the postoperative behavior should surely not be considered a recovered behavior, since it did not exist preoperatively and thus, in point of fact, could not be recovered. To say that learning subsequent to brain damage is behavioral recovery is to complicate the issue by emphasizing the outcome of a process, the attainment of some behavior, rather than the process itself, learning.

SOME POSSIBLE MECHANISMS GOVERNING RECOVERY OF FUNCTION

If the above analysis has any merit, and if it is granted that the nervous system is best considered within the conceptual framework of other biological systems, how is it possible to explain recovery of function following brain damage? Obviously, it falls somewhat short to simply say that if a specific behavior is recovered following neural insult, this behavior was never really lost. Something more concrete is needed,

but presently the data base does not allow any firm theoretical conclusions. However, there are three models that appear to have considerable merit and that are compatible with the notions of neural stability and sparing. One of these models is concerned with the reaccessing of memory engrams and has been recently formulated by Meyer (1972). A second involves the inhibition and/or facilitation of surviving neural centers by other spared neural centers (Sprague, 1966). The third is a reactivation hypothesis originally proposed by Lashley (1932, 1935) and recently revived by LeVere and Morlock (1973, 1974). It is perhaps worthwhile to conclude the present discussion with a brief encapsulation of these positions, since they subsume much of the currently available data and will most probably weigh heavily in future inquiry into recovery of function. For a full appreciation of these propositions, the reader is urged to consult the original sources.

The reaccessing notion (Meyer, 1972) emphasizes the fact that behavior depends not only on memory engrams but also the individual's ability to access these memory engrams. As neatly argued by Meyer, the disruptive effects of neural damage could be precipitated either by the destruction of the memory engram itself or, alternatively, by disrupting the ability of the individual to access the memory engram. Since, as previously outlined, it appears that specific neural mechanisms (memory engrams) are often spared following neural lesions, the accessing mechanism or mechanisms involved may be pivotal in the recovery process. Moreover, if it is accepted that the nervous system is functionally stable, then the destruction of the memory engrams supporting a particular behavior would surely result in the loss of the behavior, with little hope, if relearning is excluded, for functional recovery. However, to strictly align the reaccessing concept with the current notion of functional stability, it must be assumed that the accessing routes to memory engrams are more diffuse than the memory engrams themselves. While Meyer (1972) does not make this specific assumption, it is not a terribly unlikely assumption if one

considers the spread of evoked electrical activity following localized brain stimulation.

Another model compatible with the notions of stability and sparing has been presented by Sprague (1966). This approach is somewhat more of a systems notion than that suggested by Meyer (1972) and suggests that various behavioral functions are mediated at cortical and subcortical levels by neural centers organized into systems. Neurological insult causes dysfunction because the system is thrown out of balance by changing inhibitory/facilitory relations between the various components of the damaged system. With regard to the visual system, perhaps Sprague's (1966) comments are most succinct:

Visual attention and perception are mediated at both forebrain and midbrain levels which interact in their control of visually guided behavior. Hemianopia caused by cortical lesion is due to an imbalance of these neural centers that subserve vision, resulting in an alternation of function at the midbrain level. (p. 1547)

The final model, although obviously not the final word, is that originally suggested by Lashley (1932, 1935). Descriptively, the position may be summarized as a reactivation conceptualization where, similarly to Sprague, the mammalian nervous system is considered to be organized into a number of separate subsystems. Neural lesions cause behavioral degradation, because the individual tends to maximize its interaction with the environment and to utilize only those neurological systems functioning most appropriately. This then shifts the individual's behaviors away from those dependent on the damaged system and toward those dependent on the undamaged systems. For example, behavioral control may shift from visual cues to olfactory cues following striate neocortical ablation, with the result that visually guided behaviors would suffer. Behavioral recovery in this situation involves reactivating the damaged neural system by forcing its utilization to guide behavior. When this is accomplished, those behaviors that can be supported by the spared portions of the damaged system should return, and those behaviors that cannot be supported should not return. A specific example of

the operation of this sort of mechanism is provided by the "cross-suture" technique, which may demonstrate certain residual visual functions following early monocular deprivation in kittens (Ganz & Haffner, 1974).

SUMMARY

Whatever model of behavioral recovery is ultimately developed, it would seem clear that it must remain within a certain operational framework. The postulate of the present discussion is that this operational framework is that which is characteristic of biological systems in general and encompasses functional stability and postoperative behavior dependent on the continued operation of what is spared. Conceptualizations of behavioral recovery suggesting other vaguely specified characteristics, such as functional reorganization, should, at least for the present, be avoided. The reason for doing so is not simply a concession to parsimony but rather a respect for empirical findings.

REFERENCES

- Adametz, J. H. Rate of recovery of functioning in cats with rostral reticular lesions. An experimental study. *Journal of Neurosurgery*, 1959, 16, 85-97.
- Adams, A. D., & Forrester, J. M. The projection of the rat's visual field on the cerebral cortex. *Quarterly Journal of Experimental Physiology*, 1968, 53, 327-336.
- Bard, P., & Brooks, C. M. Localized cortical control of some postural reactions in the cat and rat together with evidence that small cortical remnants may function normally. *Association for Research in Nervous and Mental Diseases*, 1934, 13, 107-157.
- Bauer, J. H., & Cooper, R. M. Effects of posterior cortical lesions on performance of a brightness discrimination task. *Journal of Comparative and Physiological Psychology*, 1964, 58, 84-92.
- Blakemore, C., & Cooper, G. F. Development of the brain depends upon the visual environment. *Nature*, 1970, 228, 477-478.
- Bland, B. H., & Cooper, R. M. Experience and vision of the posterior neocorticate rat. *Physiology and Behavior*, 1970, 5, 211-214.
- Braun, J. J., Meyer, P. M., & Meyer, D. R. Sprouting of a brightness habit in rats following visual decortication. *Journal of Comparative and Physiological Psychology* 1966, 61, 79-82.
- Butler, R. A., Diamond, I. T., & Neff, W. D. Role of auditory cortex in discrimination of changes in frequencies. *Journal of Neurophysiology*, 1957, 20, 108-120.
- Cass, D. T., Sutton, T. J., & Mark, R. F. Competition between nerves for functional connexions with axolotl muscles. *Nature*, 1973, 243, 201-203.
- Cooper, R. M., Blochert, K. P., Gillespie, L. A., & Miller, L. J. Translucent occluders and lesions of posterior neocortex in the rat. *Physiology and Behavior*, 1972, 8, 693-697.
- Cunningham, T. J. Sprouting of optic projections after cortical lesions. *Anatomical Record*, 1972, 172, 289.
- Cunningham, T. J., & Lund, R. D. Laminar patterns in the dorsal division of the lateral geniculate nucleus of the rat. *Brain Research*, 1971, 34, 394-398.
- Dews, P. B., & Wiesel, T. N. Consequences of monocular deprivation on visual behaviour in kittens. *Journal of Physiology*, 1970, 206, 437-455.
- Finger, S., Marshak, R. A., Cohen, M., Scheff, S., Trace, R., & Niemand, D. Effects of successive and simultaneous lesions of somatosensory cortex on tactile discrimination in the rat. *Journal of Comparative and Physiological Psychology*, 1971, 77, 221-227.
- Freeman, R. D., & Thibos, L. N. Electrophysiological evidence that abnormal early visual experience can modify the human brain. *Science*, 1973, 180, 876-878.
- Ganz, L., & Haffner, M. E. Permanent perceptual and neurophysiological effects of visual deprivation in the cat. *Experimental Brain Research*, 1974, 20, 67-87.
- Goodale, M., & Cooper, R. M. Cues utilized by normal and posterior neocorticate rats in the Yerkes brightness discrimination box. *Psychonomic Science*, 1965, 3, 513-514.
- Goodman, D. C., Bogdasarian, R. S., & Horel, J. A. Axonal sprouting of ipsilateral optic tract following opposite eye removal. *Brain, Behavior and Evolution*, 1973, 8, 27-50.
- Goodman, D. C., & Horel, J. A. Sprouting of optic tract projections in the brainstem of the rat. *Journal of Comparative Neurology*, 1966, 127, 71-88.
- Grossman, S. P. *A textbook of physiological psychology*. New York: Wiley, 1967.
- Hayhow, W. R., Sefton, A., & Webb, C. Primary optic centers of the rat in relation to the terminal distribution of the crossed and uncrossed optic nerve fibers. *Journal of Comparative Neurology*, 1962, 118, 295-322.
- Hirsch, H. V. B., & Spinelli, D. N. Modification of the distribution of receptive field orientation in cats by selective visual exposure during development. *Experimental Brain Research*, 1971, 13, 509-527.
- Hubel, D. H., & Wiesel, T. N. Binocular interaction in striate cortex of kittens reared with artificial squint. *Journal of Neurophysiology*, 1965, 28, 1041-1059.

- Hubel, D. H., & Wiesel, T. N. The period of susceptibility to the physiological effects of unilateral eye closure in kittens. *Journal of Physiology*, 1970, 206, 419-436.
- Kappers, C. U. A. Further contributions on neurobiotaxis. *Journal of Comparative Neurology*, 1917, 27, 261-298.
- Lashley, K. S. Studies of cerebral function in learning. VII. A reanalysis of data on mass action in the visual cortex. *Journal of Comparative Neurology*, 1932, 54, 77-84.
- Lashley, K. S. The mechanism of vision: VIII. The projection of the retina upon the cerebral cortex of the rat. *Journal of Comparative Neurology*, 1934, 69, 57-69.
- Lashley, K. S. The mechanism of vision. XII. Nervous structures concerned in the acquisition and retention of habits based on reactions to light. *Comparative Psychology Monographs*, 1935, 11, No. 2.
- LeVere, T. E., & Morlock, G. W. The nature of visual recovery following posterior neocortical ablation in the hooded rat. *Journal of Comparative and Physiological Psychology*, 1973, 83, 62-67.
- LeVere, T. E., & Morlock, G. W. The influence of preoperative learning on the recovery of a successive brightness discrimination following posterior neocortical ablation in the hooded rat. *Bulletin of Psychonomic Society*, 1974, 4, 507-509.
- Liu, C. N., & Chambers, W. W. Intraspinal sprouting of dorsal root axons. *Archives of Neurology and Psychiatry*, 1958, 79, 46-61.
- Lourie, H., Vanasupa, P., & O'Leary, J. L. Experimental observations upon chronic progressive lesions of the brainstem tegmentum and midline thalamus. *Surgical Forum*, 1960, 10, 756-760.
- Lund, R. D., Cunningham, T. J., & Lund, J. S. Modified optic projections after unilateral eye removal in young rats. *Brain, Behavior and Evolution*, 1973, 8, 51-72.
- Lynch, G., Matthews, D. A., Mosko, S., Parks, T., & Cotman, C. Induced acetylcholine esterase-rich layer in rat dentate gyrus following entorhinal lesions. *Brain Research*, 1972, 42, 311-318.
- Mark, R. F., & Marotte, L. R. The mechanism of selective reinnervation of fish eye muscles: III. Functional, electrophysiological and anatomical analysis of recovery from section of the IIIrd and IVth nerves. *Brain Research*, 1972, 46, 131-148.
- Mark, R. F., Marotte, L. R., & Mart, P. E. The mechanism of selective reinnervation of fish eye muscles. IV. Identification of repressed synapses. *Brain Research*, 1972, 46, 149-157.
- Marotte, L. R., & Mark, R. F. The mechanism of selective reinnervation of fish eye muscle. I. Evidence from muscle function during recovery. *Brain Research*, 1970, 19, 41-51. (a)
- Marotte, L. R., & Mark, R. F. The mechanism of selective reinnervation of fish eye muscle. II. Evidence from electronmicroscopy of nerve endings. *Brain Research*, 1970, 19, 53-62. (b)
- McCouch, G. P., Austin, G. M., Liu, C. N., & Liu, C. Y. Sprouting as a cause of spasticity. *Journal of Neurophysiology*, 1958, 21, 205-216.
- Meyer, D. R. Access to engrams. *American Psychologist*, 1972, 27, 124-133.
- Meyer, D. R., & Woolsey, C. N. Effects of localized cortical destruction on auditory discriminative conditioning in cat. *Journal of Neurophysiology*, 1952, 15, 149-162.
- Meyer, P. M. Recovery of function following lesions of the subcortex and neocortex. In D. G. Stein, J. J. Rosen, & N. Butters (Eds.), *Plasticity and recovery of function in the central nervous system*. New York: Academic Press, 1974.
- Meyer, P. M., Horel, J. A., & Meyer, D. R. Effects of dl-amphetamine upon placing response in neocorticate cats. *Journal of Comparative and Physiological Psychology*, 1963, 56, 402-404.
- Mize, R. R., Wetzel, A. B., & Thompson, V. E. Contour discrimination in the rat following removal of posterior neocortex. *Physiology and Behavior*, 1971, 6, 861-867.
- Montero, V. M. Evoked responses in the rat's visual cortex to contralateral, ipsilateral and restricted photic stimulation. *Brain Research*, 1973, 53, 192-196.
- Moore, R. Y., Björklund, A., & Stenevi, U. Plastic changes in the adrenergic innervation of the rat septal area in response to denervation. *Brain Research*, 1971, 33, 13-35.
- Pettigrew, J. D., & Freeman, R. D. Visual experience without lines: Effect on developing cortical neurons. *Science*, 1973, 182, 599-601.
- Powley, T. L., & Keesey, R. E. Relationship of body weight to the lateral hypothalamic feeding syndrome. *Journal of Comparative and Physiological Psychology*, 1970, 70, 25-36.
- Raab, D. H., & Ades, H. W. Cortical and mid-brain mediation of a conditioned discrimination of acoustic intensities. *American Journal of Psychology*, 1946, 59, 59-83.
- Raisman, G. Neuronal plasticity in the septal nuclei of the adult rat. *Brain Research*, 1969, 14, 25-48.
- Ramón y Cajal, S. [*Degeneration and regeneration of the nervous system*] (Vol. 2; R. M. May, Ed. and trans.). New York: Hafner, 1959. (Originally published, 1928.)
- Rose, J. E., Malis, L. I., Kruger, L., & Baker, C. F. Effects of heavy ionizing monoenergetic particles on the cerebral cortex. *Journal of Comparative Neurology*, 1960, 115, 243-295.
- Rosenzweig, M. R. Discrimination of auditory intensities in the cat. *American Journal of Psychology*, 1946, 59, 127-136.
- Rosner, B. S. Brain functions. In D. Mussen & M. Rosenzweig (Eds.), *Annual Review of Psychology*. Palo Alto, Calif: Annual Reviews, 1970.

- Rosner, B. S. Recovery of function and localization of function in historical perspective. In D. G. Stein, J. J. Rosen, & N. Butters (Eds.), *Plasticity and recovery of function in the central nervous system*. New York: Academic Press, 1974.
- Schneider, G. E. Mechanisms of functional recovery following lesions of visual cortex or superior colliculus in neonate and adult hamsters. *Brain, Behavior and Evolution*, 1970, 3, 293-323.
- Schneider, G. E. Early lesions of superior colliculus: Factors affecting the formation of abnormal retinal projections. *Brain, Behavior and Evolution*, 1973, 8, 73-109.
- Spear, P. D., & Braun, J. J. Nonequivalence of normal and posteriorly neocorticated rats on two brightness discrimination problems. *Journal of Comparative and Physiological Psychology*, 1969, 67, 235-239.
- Sperry, R. W. Nature of functional recovery following regeneration of the oculomotor nerve in amphibians. *Anatomical Record*, 1947, 97, 293-315.
- Sperry, R. W. Myoptic specificity in teleost motoneurons. *Journal of Comparative Neurology*, 1950, 93, 277-288.
- Sperry, R. W. Problems in the biochemical specifications of neurons. In H. Waelsch (Ed.), *Biochemistry of the developing nervous system*. New York: Academic Press, 1955.
- Sperry, R. W. Physiological plasticity and the brain circuit theory. In H. F. Harlow & C. N. Woolsey (Eds.), *Biological and biochemical basis of behavior*. Madison: University of Wisconsin Press, 1958.
- Sperry, R. W., & Hibbard, E. Regulative factors in the orderly growth of retinotectal connexions. In G. E. W. Wolstenholme & M. O'Connor (Eds.), *Growth of the nervous system*. Boston: Little, Brown, 1968.
- Sprague, J. M. Interaction of cortex and superior colliculus in mediation of visually guided behavior in the cat. *Science*, 1966, 153, 1544-1547.
- Stenevi, U., Björklund, A., & Moore, R. Y. Growth of intact central adrenergic axons in the denervated lateral geniculate body. *Experimental Neurology*, 1972, 35, 290-299.
- Teitelbaum, P., Cheng, M., & Rozin, P. Development of feeding parallels its recovery after hypothalamic damage. *Journal of Comparative and Physiological Psychology*, 1969, 67, 430-441.
- Teitelbaum, P., & Epstein, A. The lateral hypothalamic syndrome. *Psychological Review*, 1962, 69, 74-90.
- Wall, P. D., & Egger, M. D. Formation of new connexions in adult rat brains after partial deafferentation. *Nature*, 1971, 232, 542-545.
- Wiesel, T. E., & Hubel, D. H. Single-cell responses in striate cortex of kittens deprived of vision in one eye. *Journal of Neurophysiology*, 1963, 26, 1003-1017.
- Weiss, P. Selectivity controlling the central-peripheral relations in the nervous system. *Biological Review*, 1936, 11, 494-531.
- Winans, S. S. Visual form discrimination after removal of visual cortex in cats. *Science*, 1967, 158, 944-946.

(Received December 19, 1974)

Time-Dependent Variations in Aversively Motivated Behaviors: Nonassociative Effects of Cholinergic and Catecholaminergic Activity

Hymie Anisman
Carleton University, Ottawa, Canada

The role of acetylcholine and central catecholamines in modulating aversively motivated behaviors, and behavior following exposure to uncontrollable stressors, is evaluated. It is suggested that in the presence of adequate associative processes, nonassociative factors, mediated by stress-induced neurochemical changes, determine avoidance response rate. Moreover, in addition to the existence of a balanced state between excitatory catecholamine and inhibitory cholinergic systems, it is posited that these neurochemical systems may be mutually regulatory. Excessive stimulation of one system may induce a compensatory antagonistic rebound in the complementary system, thereby maintaining neurochemical homeostasis. Owing to time-dependent variations in neurotransmitter activity, temporal variations in performance may occur following initial exposure to aversive stimulation. Alterations in neurochemical activity that affect nonassociative processes have predictable effects on time-dependent variations in avoidance performance. The model is extended to deal with other stress related phenomena, for example, helplessness, depression, and ulceration.

It has become increasingly apparent that aversively motivated behaviors, as well as other behaviors involving response inhibition (e.g., habituation, responsivity to nonreinforced stimuli, shock-induced response suppression), are modulated largely by the activities of acetylcholine, catecholamines, and indoleamines (cf. reviews by Bignami, Rosic, Michalek, Milosevic, & Gatti, 1975; Carlton, 1963, 1969; Carlton & Markiewicz, 1971; Rech & Stolk, 1970; Stein & Wise, 1974). Moreover, although modification of levels of any one of these transmitters may have profound effects on behavior quite independently of other transmitter systems, ample data are available indicating that these systems act in a reciprocally coordinated fashion (Carlton, 1963; Chase, 1974; Costall & Naylor, 1974; Gellhorn, 1957). Essentially,

the suggestion has been that the catecholamines, whose behavioral concomitants involve excitation, act antagonistically toward the cholinergic and serotonergic systems that exert an inhibitory influence upon ongoing behaviors.

To date, the notion of reciprocally antagonistic systems has fared well in accounting for phenomena such as drug interactions (e.g., Carlton, 1963; McGeer, Grewaal, & McGeer, 1974; Swonger & Rech, 1972), ontogenic variations in inhibitory processes (Campbell, Lytle, & Fibiger, 1969; Campbell & Mabry, 1973), parkinsonism (Bartholini, Stadler, & Lloyd, 1973), stereotypic behaviors (Arnfred & Randrup, 1968; Janowsky, El-Yousef, Davis, & Sekerke, 1972), and factors involved in habituation (Carlton, 1968, 1969). Surprisingly though, a paucity of data exists regarding the role one neurotransmitter system may have in modulating the effects, and indeed in altering the actual levels, of a second system and how such interactions might ultimately affect performance levels. There is reason to believe that when a particular transmitter system is excited excessively, an antagonistic rebound may be induced in a second system,

Supported by Grant A9845 from the National Research Council of Canada. Appreciation is extended to Giorgio Bignami, Peter Carlton, Douglas Wahlsten, Gary Waller, Robert Lahue, Gary Remington, Larry Kokkinidis, Robert Hoffman and Peter Fried for critical comments on an earlier draft of this paper.

Requests for reprints should be sent to Hymie Anisman, Department of Psychology, Carleton University, Ottawa, Ontario, K1S 5B6, Canada.

thereby maintaining balance between excitatory and inhibitory processes (Agid, Guyenet, Javoy, Beaujouan, & Glowinski, 1974; Bartholini et al., 1973; Gellhorn, 1957; Manto, 1967; McGeer et al., 1974; Richardson, 1974; Sethy & Van Woert, 1974).

In the present paper the balance between transmitter systems and the possibility of their reciprocal innervation are evaluated with respect to effects on aversively motivated behaviors in general and to the effects of prior aversive stimulation on subsequent behaviors in particular. In the sections that follow, it is shown that performance in aversively motivated tasks, although invariably dependent on associative processes, are modified by nonassociative events. Because the temporal effects of stress are not necessarily monotonic, an extended discussion of time-dependent variations in performance is presented. Finally, neurochemical concomitants of performance following the inception of a stressor are discussed. These neurochemical events are proposed to mediate nonassociative events that affect avoidance performance.

Time-Dependent Changes in Aversively Motivated Behaviors

It has been demonstrated repeatedly that following exposure to signaled inescapable shock or only a few escape/avoidance trials, subsequent active avoidance performance progresses more slowly after an intermediate retention interval (1-6 hours) than after a short (less than 10 minutes) or long (24 hours) interval (Brush, 1971; Kamin, 1963). The U-shaped performance function is of some interest, because retention functions for most associative learning situations are monotonic and negatively decelerating, thus making the avoidance retention function somewhat paradoxical in nature.

While several hypotheses have been formulated to elucidate the mechanisms subserving the time-dependent changes in behavior, none of these positions adequately account for the available data (cf. Anisman, 1973a; Barrett, Leith, & Ray, 1973; for a review of most of the current hypotheses, see Brush, 1971). Though critiques of some

of these hypotheses are advanced later in this paper, it is important to note that the alternative positions expounded by Brush (1971), involving availability of corticosterone and coping abilities, and by Klein and Spear (1970a, 1970b), dealing with memory retrieval, are not necessarily contrary to the position advanced in the present paper. Essentially, the thesis of the current proposal is that initial exposure to stressful stimulation has both associative and non-associative consequences, which interact to determine the rate at which a subsequent avoidance response will be acquired. Whereas changes in associative processes are thought to be minimal over the initial 24-hour period, nonassociative changes elicited by stress are not monotonic. Ultimately, in the presence of adequate associative abilities, nonassociative changes may determine the form of the retention function.

To be more explicit, in order for an active avoidance response to be acquired, several components are necessary. A drive state must be established, defined in terms of Pavlovian processes (cf. Rescorla & Solomon, 1967) or expectancy of shock (Bolles, 1970, 1971); the animal must learn *when* to initiate an active response, appropriately direct this response (Anisman & Wahlsten, 1974), and be able to *initiate* the active locomotor response (cf. Anisman & Waller, 1973a). While it is clear that associative factors ultimately set rate limitations on avoidance acquisition, the role of nonassociative factors in determining response rate only recently has become a major source of debate with respect to aversively motivated behaviors (cf. Bolles, 1970, 1971; Seligman, 1970). In the typical active avoidance situation, it has been observed that the rate of acquiring the avoidance response was inversely related to the strength of an incompatible freezing response produced by shock and directly related to the rate at which this freezing response could be suppressed (Anisman & Waller, 1972, 1973a; Bolles, 1970).

Thus, it has been postulated that in the presence of adequate associative abilities, manipulations that increase freezing behavior elicited by shock will disrupt performance, whereas those that decrease inhibitory ten-

dencies will augment active avoidance behaviors. In the absence of adequate associative abilities, it is unlikely that reduction in nonassociative effects of shock will alter avoidance behavior substantially (Anisman, 1975, in press; Anisman, Wahlsten, & Kokkinidis, in press). With respect to the time-dependent changes in avoidance, the suggestion has been that at an intermediate interval, stimulus and response contingencies previously established are present. However, animals are less able to initiate adequately the active response necessary for successful avoidance, and as a result performance at this interval is inferior to that seen at brief or long intervals, when the locomotor response can be initiated with somewhat less difficulty (Anisman, 1973a; Anisman & Kokkinidis, 1974; Anisman & Waller, 1971; Barrett, Leith, & Ray, 1971a, 1972; Pinel & Mucha, 1973a, 1973b).

Considerable support has accrued for a hypothesis based on the nonassociative consequences of shock. This support has been derived from studies aimed at examining response-repertoire changes following exposure to noncontingent shock, as well as from studies involving response-shock contingencies (active and passive avoidance). Specifically, following exposure to noncontingent shock, behavior is characterized by levels of activity that exceed baseline (Kulmar, 1970; Pinel, Corcoran, & Malsbury, 1971; Pinel & Mucha, 1973a, 1973b). The period of excitation is transient, subsiding within 10 minutes and falling below baseline. While activity during the phase characterized by response suppression is inversely related to shock intensity, the initial hyperactivity is directly related to shock intensity (Anisman & Waller, 1973b). Within 24 hours after shock exposure activity levels increase and approach baseline. The actual rate at which activity increases is dependent on several factors, such as shock intensity (Anisman & Waller, 1973b) and the strain of animal employed (Anisman, 1973a).

In experiments involving specific response-shock contingencies, such as conditioned emotional response (CER) and passive avoidance tasks, it was observed that at

short periods after initial shock there is a phase of hypermotility, which results in a weak CER and poor avoidance performance. Between the brief and intermediate interval there is an increased CER and improved passive avoidance performance (Anisman & Waller, 1973b; Geller, Jarvik, & Robustelli, 1970; Irwin, Banuazizi, Kalsner, & Curtis, 1968; Kumar, 1970; McMichael, 1966; Robustelli, Geller, & Jarvik, 1970; Tarpy, 1966; Zammit-Montebello, Black, Marquis, & Suboski, 1969). Evidently, although the pairing of a stressor with an explicit cue has profound associative consequences, the initial shock-induced excitation and subsequent response suppression may modify performance. That is, during the brief interval after shock, the response bias created is incompatible with passive avoidance, whereas at the intermediate interval, the shock-induced response bias favors passive avoidance and response suppression.

Owing to the nature of the passive avoidance task, it is difficult to divorce associative and nonassociative factors that contribute to the ultimate levels of performance. However, when considered together with active avoidance performance, the influence of nonassociative factors becomes apparent. For example, in a series of experiments conducted by Barrett and his associates (Barrett, Leith, & Ray, 1971b, 1972; Steranka & Barrett, 1973, 1974), it was observed that following initial stress both activity levels and Y-maze avoidance performance resembled U-shaped functions and that activity levels correlated directly with avoidance response rate. In contrast, discrimination performance in the Y-maze avoidance/escape task did not vary over the various retention intervals. Evidently, animals "know where to run" (i.e., the memory of the safe and dangerous locales is intact), but at the intermediate retention interval animals suffer from an inability to initiate the active locomotor response necessary for successful avoidance.

In support of such an interpretation, Anisman (1973a) observed that among Holtzman rats a U-shaped performance function was observed in a shuttle task following initial exposure to signaled inescap-

able shock. However, if animals were tested in a one-way avoidance task the U-shaped performance function was not observed. It seems that under those conditions in which freezing cannot be suppressed readily, as in the shuttle task, differences in freezing behavior effectively modulate avoidance response rate. In contrast, when freezing behavior can be suppressed in relatively few trials, as in the case of the one-way avoidance task, the nonassociative effects elicited by initial shock exposure have only very minor effects on avoidance performance. Together with the data for passive avoidance, it seems that at the intermediate interval, the high levels of response suppression favor inhibitory tasks and may interfere with an active avoidance task. In contrast, at the brief interval, the presence of response excitation favors active responding but may hinder passive avoidance. Additional evidence in support of this contention is presented later in this paper (see section on pharmacological manipulations).

It is clear from the preceding discussion that one of the primary assumptions of a nonassociative explanation is that the U-shaped function is a response to shock *per se*, rather than cue shock pairings. Thus at the brief and intermediate intervals the excitation and inhibition of activity, respectively, would be expected to modify behavior in a predictable manner. With the attenuation of response suppression at the long interval, response rate would be primarily an indicant of associative processes and species-typical defensive behaviors. As indicated earlier, whether variations in response inhibition will be reflected in response change (e.g., improved active avoidance) depends on a variety of associative factors. Brush (1966) observed that unlike signaled shock, initial exposure to unsignaled shock does not promote a U-shaped function. Presumably, if the stimulus-shock contingencies are not well established during initial training, then potential differences in performance due to nonassociative factors may not be manifested. For example, under conditions in which animals receive exposure to shock in a nonilluminated compartment, a conflicting response tendency will be pro-

duced when animals subsequently are required to avoid shock by running from an illuminated (i.e., by the conditioned stimulus) to a nonilluminated chamber (Anisman, 1973b). Accordingly, subsequent performance at all intervals initially will be retarded, and differences in freezing behavior at the various intervals will not be detectable.

It should be considered further that when the contingencies of the avoidance task are ambiguous even under the best of conditions (e.g., as in shuttle avoidance), the low levels of performance typically observed may add to the obfuscation of the effects of the non-associative factors. Moreover, since tasks such as shuttle avoidance appear to be particularly dependent on feedback stimuli (cf. Anisman & Waller, 1973a; Bolles, 1971), the initial establishment of the stimulus-response contingencies take on all the more importance in the derivation of the U-shaped function. Predictably, if avoidance testing is carried out in a one-way task, which typically is not plagued by low levels of performance, and is less dependent on explicit feedback stimuli than the shuttle task, then a U-shaped retention function is observed even when an explicit conditioned stimulus is not paired with shock during initial training (Anisman & Waller, 1971; Bauer, 1973; de Toledo & Black, 1970). More recently, Riley, Dunlap, and Hughes (1975) also reported that the U-shaped performance function was observed in a one-way task following exposure to inescapable unsignaled shock in a compartment radically different from that used during avoidance training. Together with the finding that the U-shaped function was observed under these conditions even after a single inescapable shock, the nonassociative explanation becomes more tenable.

The interaction between shock and cue-shock pairings in determining the U-shaped function has been elegantly shown in experiments carried out by Pinel and Mucha (1973a, 1973b) and by Barrett et al. (1971b). Specifically, Pinel and Mucha (1973a) observed that the administration of a second shock during the period when activity was low resulted in an initial excita-

tion and subsequent suppression of locomotor activity comparable to that seen after the initial shock session. In effect, the suggestion made by these investigators was that time since the last shock and not time since initial training determined nonassociative processes. Similarly, Barrett et al. (1971b) reported that if animals initially received avoidance training involving clearly discriminable cues, thus establishing the contingencies of the avoidance task, and 24 hours later were subjected to unsignaled inescapable shock, subsequent avoidance performance at various intervals following inescapable shock resembled the typical U-shaped function. This particular experiment illustrates again that the U-shaped function is dependent on time since the most recent stress and not on time since original learning. Moreover, if the task was clearly discriminable during original training, then a U-shaped function was observed following unsignaled exposure to shock.

These studies, taken together, seem to suggest that the simple presence of nonassociative changes following shock will not insure the occurrence of a U-shaped function. The rate-limiting step must be one of adequate associative properties being present. Inasmuch as the interaction between nonassociative factors and associative processes determines performance, and nonassociative processes are time dependent, this preparation constitutes an ideal situation for divorcing the role of biochemical events that mediate nonassociative processes from those situations in which biochemical changes occur and response contingencies are varied.

Time-Dependent Variations in Neurochemical Activity and Avoidance

Clearly, initial exposure to shock has nonassociative consequences that influence avoidance behavior. What remains to be understood is what factors mediate the time-dependent variations of these nonassociative effects. At an intuitive level it would appear most parsimonious to assume that freezing behavior, and its underlying mechanism, should be maximal soon after stress and should dissipate thereafter. However, if such a model is accepted, a second factor

must be invoked to account for the absence of freezing behavior and the high level of performance observed at the brief interval. One possibility is that shock elicits transient biochemical changes that retard or antagonize the course of response inhibition that might otherwise occur. In effect, it is assumed that mechanisms that produce response suppression are at least *partially* operative, but the stimulation of a second system masks these suppressive effects. At this point, further biochemical mechanisms can be introduced. These are assumed to be more sluggish in nature and may be a consequence of the initial biochemical change. The maximal effects of these changes are thought to occur at an intermediate interval, and these tend to exacerbate response inhibition. Thus, the U-shaped function could well be due to interaction between the initial biochemically induced behavioral excitation and the somewhat postponed behavioral inhibition elicited by further chemical alteration. At relatively long intervals following the initial stress, chemical levels should approach a base state, and consequently, performance at this interval would more closely approximate the uncontaminated associative effects of the initial training.

Recently, the nonassociative effects of shock have been attributed to variations of several neurotransmitters (Anisman, 1973a; Anisman & Kokkinidis, 1974; Barrett, Leith, & Ray, 1971b; Manto, 1967; Thompson & Neilson, 1972; Brett, Note 1). Together, these studies have yielded an integrated model involving the interaction and mutual innervation between several neurotransmitter systems. Specifically, as depicted in Figure 1, exposure to shock is thought to elicit a transient catecholamine (dopamine and norepinephrine) action, followed by a cholinergic rebound together with subsequent decreased catecholamine activity (cf. Anisman, 1973a; Anisman & Kokkinidis, 1974; Barrett et al., 1971b; Manto, 1967; Thompson & Neilson, 1972; Brett, Note 1). The initial catecholamine action is thought to result in behavioral excitation (Costa & Groppetti, 1970; Fuxe & Ungerstedt, 1970) and consequently proficient avoidance be-

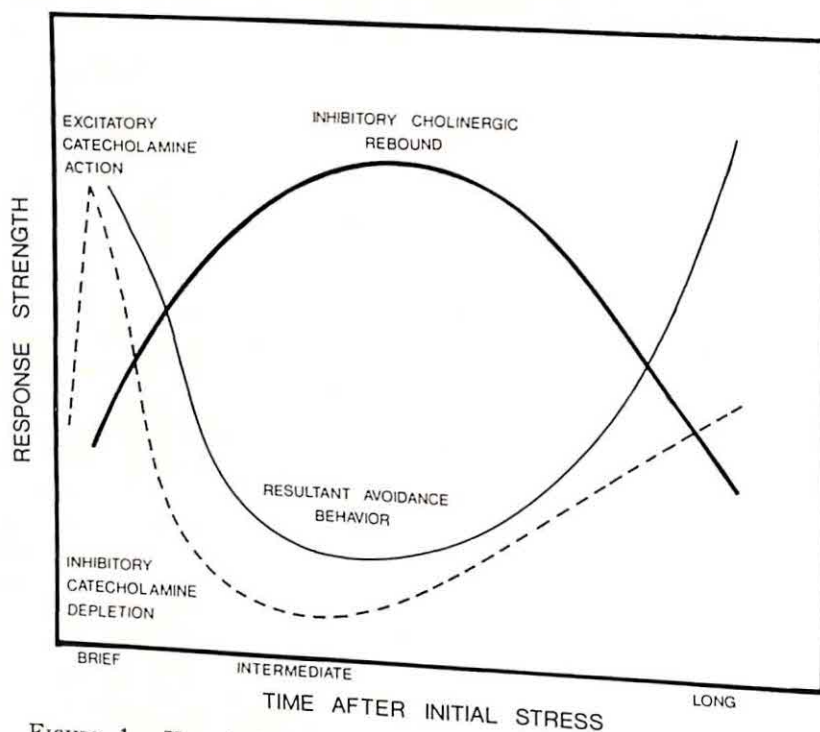


FIGURE 1. Hypothetical curves for catecholaminergic and cholinergic activity as well as the resultant avoidance curve. Note that immediately following the inception of stress, catecholamine activity shows a transient increase and then falls below baseline. Cholinergic action increases steadily, reaching maximum at the intermediate interval.

havior (Breese, Cooper, & Smith, 1973; Rech & Stolk, 1970). The catecholamine action in turn results in a compensatory cholinergic rebound, which produces behavioral inhibition (cf. Carlton, 1969) and poor avoidance performance (Anisman, 1973a; Anisman & Kokkinidis, 1974; Bignami & Gatti, 1966). Moreover, following the initial catecholamine excitation, catecholamine activity declines, thus exacerbating the inhibitory effects of the cholinergic reaction. As seen in Figure 1, unlike the transient excitatory catecholamine action, the cholinergic reaction is more sluggish and longer lasting. During those periods when the balance between the two systems favors cholinergic action, response initiation is inhibited and avoidance progresses slowly, whereas the converse is true of those periods when the balance rests in favor of catecholamine activity.

The data in support of such a model are derived from several sources. In the sec-

tions that follow pertinent data are presented indicating that (a) balance and reciprocal innervation between neurotransmitter systems may occur in both the central and peripheral nervous system, (b) exposure to a stressor has profound effects on the activity of several neurotransmitters, and (c) pharmacological manipulations have predictable effects on avoidance in general and on the U-shaped function in particular. Moreover, these pharmacological effects are due to modification of nonassociative as opposed to associative factors.

Balance and Reciprocal Innervation of Neurotransmitters

The notion of reciprocally activated systems is not unique in the literature. For example, in a thorough documentation of sympathetic and parasympathetic activity, Gellhorn (1957) indicated that not only does balance exist between the adrenergic and cholinergic systems, but excitation of

one system may result in the activation of the complementary system. Specifically, sympathetic stimulation may lead to a decreased threshold, as well as an increased reaction to parasympathetic stimulation. Moreover, the parasympathetic discharge may summate with further parasympathetic stimulation or be suppressed by concurrent sympathetic stimulation. Gellhorn further reported that stimulation of the hypothalamus could elicit sympathetic signs; however, upon termination of the hypothalamic stimulation, parasympathetic responsivity was observed. Similarly, if epinephrine was administered prior to stimulation of the posterior hypothalamus, sympathetic responses ordinarily elicited by the hypothalamic stimulation were eliminated. On the basis of these observations, it was suggested that the parasympathetic reaction is directly related to the magnitude (intensity, frequency, or duration) of the preceding sympathetic stimulation. The function of the reciprocity is to maintain neurochemical homeostasis between the antagonistic portions of the autonomic system. Moreover, the rebound presumably is mediated by baroreceptors of the carotid sinus and aortic arch, since transection of nerves connecting the baroreceptors with the medulla prevents the occurrence of rebound effects. While such a mechanism could possibly account for sympathetic and parasympathetic balance, it is unlikely that this same mechanism accounts for aversively motivated behaviors, since peripherally acting agents typically have not been observed to produce large changes in avoidance behavior.

Recently, it has been observed that sympathetic excitation may result in a central cholinergic rebound similar to that stipulated by Gellhorn (1957) regarding parasympathetic activity, and in addition, the balance resulting from the central innervation may modulate behaviors such as general activity and avoidance behavior. Specifically, Manto (1967) observed that 1 hour after injection of epinephrine there is a decline in general activity. The decline in activity was reversed with scopolamine but not by the peripheral analogue methylscopolamine, suggesting that peripheral adrenergic

stimulation may result in an inhibitory central cholinergic rebound or that parasympathetic responsivity is mediated by central mechanisms. Remington and Anisman (1974) also have observed that the disruptive effects of epinephrine on avoidance behavior could be reversed by scopolamine. Inasmuch as phentolamine, an alpha receptor antagonist of the adrenergic system, attenuated the epinephrine-induced decrement in activity, Manto (1968) suggested that stimulation of alpha receptors, as well as beta receptors as suggested by Gellhorn, were necessary for the rebound effect to occur. This particular notion was supported further by the finding that epinephrine, which excites both alpha and beta receptors, was more effective in the induction of a cholinergic rebound than norepinephrine, which excites alpha receptors.

In a pseudoconditioning barrier-crossing task, Izquierdo (1974) observed that performance was enhanced by amphetamine. Adrenergic blocking drugs, administered in doses that had no effects on their own, successfully antagonized the amphetamine effects. Conversely, the peripheral anticholinergic, methylatropine, although producing negligible changes in behavior, potentiated the effects of amphetamine. Accordingly, it was suggested that although amphetamine exerts an excitatory effect on behavior, it also induces a parasympathetic rebound that detracts from the stimulant properties of amphetamine. Thus, with cholinergic blockade the inhibitory rebound is dampened and behavioral excitation is augmented. Parenthetically, the position that amphetamine exerts both excitatory adrenergic and inhibitory cholinergic effects may be useful in accounting for the synergistic action of amphetamine and anticholinergics (e.g., Carlton, 1961a, 1961b, 1963; Carlton & Didamo, 1961). That is, scopolamine or atropine in doses that do not produce behavioral disinhibition may be sufficient to antagonize the inhibition produced by amphetamine, thereby eliciting increased activity.

Hitherto unpublished data collected in our laboratory (Anisman & Kokkinidis, Note 2) have yielded results similar to those reported by Izquierdo (1974). Specifically, although

large doses of d-amphetamine (5 mg/kg) produce locomotor excitation in mice, lower doses (.5 or 1.0 mg/kg) induce behavioral suppression. The peripherally acting agent, p-hydroxyamphetamine, administered in equimolar doses did not affect activity, thus suggesting that both the excitatory and inhibitory effects of d-amphetamine were mediated at a central level. If animals were pretreated with scopolamine, the depressed behavior produced by d-amphetamine was not observed, whereas the peripheral cholinergic agent, methylscopolamine, did not affect behavior. Thus it seems that the inhibitory effects of d-amphetamine are mediated by central activity and possibly involve cholinergic action. It is likely that in addition to resulting in increased release and synthesis of catecholamines, a cholinergic rebound producing behavioral suppression was also elicited (e.g., Mandell & Knapp, 1971). With high doses of d-amphetamine the balance between catecholamine and acetylcholine activity rests in favor of the catecholamine system, thus producing behavioral excitation. With low doses of d-amphetamine, the balance rests in favor of central cholinergic activity and consequently is manifested by behavioral inhibition.

A similar model may possibly hold for some of the paradoxical behavioral effects elicited by L-dopa. Specifically, it is well known that L-dopa in high doses, when administered with an extracerebral decarboxylase inhibitor, results in increased levels of activity. Presumably, this is a result of the increased level of dopamine (Carlsson, 1970; Fuxe & Ungerstedt, 1970). Paradoxically, however, low doses of L-dopa result in the retardation of activity (Butcher & Engel, 1969). While peripheral factors may contribute to the decline in activity, it is clear that central mechanisms are also involved (Stromberg, 1970). It is likely that increases in dopamine result in a cholinergic rebound (e.g., Sethy & Van Woert, 1973). With high doses of L-dopa the balance between the two systems rests on the dopaminergic side and results in stimulation of behavior. With low doses the balance rests in favor of cholinergic action, and thus decreased activity is observed.

In the investigation of yet another behavioral phenomenon, namely that of stereotypic behaviors elicited by catecholamine stimulant drugs, it has been observed repeatedly that administration of cholinergic agonists and anticholinesterase agents eliminate or reduce stereotypy, whereas antimuscarinics augment stereotypy (Arnfred & Randrup, 1968; Costall & Naylor, 1972; Fjalland & Moller-Nielsen, 1974; Janowsky et al., 1972; Pedersen, 1967; Scheel-Kruger, 1970). Since stereotypy appears to be mediated largely by dopamine (Randrup & Munkvad, 1968, 1970), it has been posited that reciprocal innervation may exist between the dopaminergic and cholinergic systems. Further to this notion, it has been reported that inhibition of dopamine synthesis reduces the effects of anticholinergics (Thornburg & Moore, 1973), while dopamine stimulants are potentiated with anticholinergics (Arnfred & Randrup, 1968). Costall and Olley (1971) have shown that cataleptic behavior induced by the cholinergic agent, arecoline, can be potentiated by lesion in the caudate putamen, and reduced by lesion of the substantia nigra. It is likely that the substantia nigra is the site of action of the cholinergic drug, and feedback exists between this site and the striatum via the dopaminergic nigrostriatal system. Thus, excitation of one transmitter may cause rebound effects in a second transmitter system. Indeed, it has been reported (Ahtee & Kääriäinen, 1974) that catalepsy induced by pilocarpine was accompanied by increased homovanillic acid, the metabolite of dopamine, without an increase in brain dopamine per se, suggesting that the cholinergic stimulant produces increased metabolism of dopamine. It is noteworthy that these authors indicated that although the cholinergic and dopaminergic systems are antagonistic and can compensate for the activity of one another, this is true only to a certain extent.

Given the available data, it appears that adrenergic stimulation may promote a cholinergic rebound and thus modify behavior. Fibiger, Lynch, and Cooper (1971) have reported additional data suggesting that cholinergic stimulation may also result in an adrenergic rebound. They observed that al-

though pilocarpine, a cholinomimetic, initially induces behavioral inhibition, a subsequent phase of behavioral excitation also occurred. While pretreatment with scopolamine eliminated both the inhibitory and excitatory effects of pilocarpine, scopolamine administered with the onset of the hyperactive phase potentiated the excitation.

It is clear from the psychopharmacological literature that reciprocity between systems may occur. Only recently, however, have biochemical data become available that have demonstrated that systems are in fact mutually regulating. Specifically, it has been observed that enzymes associated with acetylcholine (ACh) are increased following injection of catecholamine stimulants such as norepinephrine (Ho, Singer, & Gershon, 1971), L-dopa (Sethy & Van Woert, 1973) and d-amphetamine (Mandel & Knapp, 1971). Decreasing levels of catecholamines via 6-hydroxydopamine, on the other hand, results in decreased activity of enzymes associated with ACh (Benton, Moffatt, Rick, & Taberner, 1975). Recently it has been suggested (Agid et al., 1974; Ladinsky, Consolo, Bianchi, Saminin, & Ghezzi, 1975; Sethy & Van Woert, 1974) that in the striatum, dopamine exerts a tonic inhibitory role on ACh activity. Dopamine stimulating agents increase ACh levels, presumably by decreasing the firing rate of ACh neurons. Conversely, dopamine blocking agents decrease ACh levels via an increase in the release of ACh. As might be predicted, the increased release of ACh by dopamine blocking drugs is reversed by dopamine stimulants (Bartholini et al., 1973; Stadler, Lloyd, & Bartholini, 1974; Stadler, Lloyd, Gadea-Ciria, & Bartholini, 1973). However Lloyd, Stadler, and Bartholini (1973) have indicated that although a balance of this nature exists in the striatum, this does not hold true for limbic structures. This notwithstanding, it is apparent that reciprocal innervation between systems does exist. Moreover, these data suggest that some behaviors, such as tremorigenic effects due to low dopamine levels, might well involve acetylcholine as well.

From the side of manipulating acetylcholine levels and gauging the change in dopa-

mine, it has been observed (Javoy, Agid, Bouvet, & Glowinski, 1974) that carbachol infusion into the substantia nigra reduced the utilization of newly formed dopamine in the neostriatum, but had no effect on synthesis. Atropine infusion, on the other hand, stimulated synthesis and utilization of dopamine. Hitzemann, Loh, and Domino (1972) reported that scopolamine treatment decreased the accumulation of ^{14}C -dopamine and ^{14}C -norepinephrine in the mouse brain after administration of either ^{14}C -tyrosine or ^{14}C -dopa. Similarly, several investigators have reported that anticholinergics antagonize increases in homovanillic acid induced by neuroleptics (Anden, 1972; Bowers & Roth, 1972; O'Keefe, Sharman, & Vogt, 1970) and also induced modification of dopamine turnover (Anden & Bedard, 1971; Corrodi, Fuxe, & Lidbrink, 1972). Finally, Herman, Kmiecik-Kolada, Slominska-Zurek, and Szkilnik (1972) observed that intravenous injection of ACh decreased levels of norepinephrine in cortex, mesencephalon, pons, and medulla, while Kazic (1973) reported that intravenous administration of physostigmine markedly increased synthesis and turnover of norepinephrine in the hypothalamus. Taken together, these data indicate that modification of either acetylcholine or dopamine, *inter alia*, will also result in a change in transmitter level in the complementary system. It seems that not only do balances exist, but reciprocal activation occurs and probably serves in a homeostatic capacity.

Effects of Stress on Neurotransmitter Activity

The position of the present paper is that modification of neurotransmitters may account for response changes following exposure to an initial stress situation. That is, initial stress may result in increased catecholamine activity, which in turn produces a central cholinergic rebound, the maximal effects of which do not occur until 1 hour after termination of the stress. Moreover, following the initial efflux of catecholamines, there is a large diminution of catecholamine action. Thus, soon after shock a state of motor hyperactivity exists, which favors ac-

tive avoidance and is incompatible with response suppression. At an intermediate interval the cholinergic dominance produces behavioral inhibition and consequently poor active avoidance performance. The involvement of the catecholamine and cholinergic systems is supported to a great extent by available data regarding the effects of stress on chemical activity.

It is well documented that stressors such as footshock, cold swim, and restraint have profound effects on the release and biosynthesis of several neurotransmitters (for reviews, see Barchas, Ciaranello, Stolk, Keith, Brodie, & Hamburg, 1972; Karczmar, Scudder, & Richardson, 1973; Weiss, Glazer, & Pohorecky, in press; Welch & Welch, 1970). Specifically, it has been reported that with the inception of relatively low levels of footshock there is a rapid increase in brain catecholamines (cf. Breitner, Picchioni, & Chin, 1963; Kato, Gozsy, Roy, & Groh, 1967; Welch & Welch, 1968, 1970). The apparent source for the increase in these amines appears to be due to inhibition of monoamine oxidase, the enzyme that degrades these amines intracellularly. Support for such a contention has come from the finding that stress increases O-methylated derivatives of catecholamines with negligible alterations in their oxidatively deaminated catabolites (Bliss, Ailion, & Zwanziger, 1968; Carlsson, 1966; Javoy, Thierry, Kety, & Glowinski, 1968; Thierry, Javoy, Glowinski, & Kety, 1968). With sufficiently intense stress, on the other hand, reduced levels of norepinephrine have been noted (Bliss et al., 1968; Bliss & Zwanziger, 1966; Maynert & Levi, 1964).

Thierry (1973) has suggested that with the inception of a relatively mild stress, newly synthesized norepinephrine in the "functional storage pool" is utilized, whereas severe stress may lead to utilization of amines stored in the "main compartment" as well, thus leading to greater decline of total norepinephrine. In any event, following administration of a strong stressor, decline in norepinephrine becomes more accentuated 2-4 hours after application of the stressor (Barchas & Freedman, 1963; Stolk, Conner, Levine, & Barchas, 1974; Stone,

1970; see also Rosecrans, 1969) and has been known to require as long as 1-6 hours before the level of norepinephrine is replenished (Barchas & Freedman, 1963; Maynert & Levi, 1964; Paulsen & Hess, 1963). The sustained reduction in norepinephrine levels could come about when neuronal activity, probably owing to activity of cells in the locus coeruleus (Korf, Aghajanian, & Roth, 1973), increases to the extent that amine release and utilization exceeds total biosynthesis (cf. Moore & Lariviere, 1963; Rosecrans, 1969; Welch & Welch, 1970). In effect, then, with the inception of a severe stressor, norepinephrine is released from the primary storage granules in large amounts, which initially may lead to behavioral excitation. However, when norepinephrine utilization exceeds synthesis, the total norepinephrine available is diminished, consequently resulting in behavioral inhibition. Weiss et al. (in press) have indicated that not only are norepinephrine levels decreased in stressed animals, but also reuptake of norepinephrine is increased, thereby further reducing the amount of norepinephrine which might impinge on the postsynaptic receptor.

While stress clearly alters norepinephrine synthesis and release, it appears that the extent of norepinephrine change may be modified by the organism's ability to cope with the stressor. Weiss, Stone, and Harrell (1970) and Weiss, Pohorecky, Dorros, Williams, Emmel, Whittlesey, and Case (Note 3) have observed that when animals can escape or avoid shock, norepinephrine levels do not decline, whereas yoked control animals show significant reduction in whole brain norepinephrine as well as in hypothalamus, brain stem, and telencephalon. Weiss et al. (in press) also indicated that when animals can cope with the stress, either through active or passive responding, norepinephrine levels may in fact increase, as opposed to the typical decrease of yoked control animals. Along these same lines, Stolk et al. (1974) observed that in rats given electric shock, norepinephrine levels were lower 3 hours after stress, relative to norepinephrine levels in nonshocked animals. However, if rats were shocked in pairs, thus eliciting fighting, norepinephrine reduction

was not observed. It may well be that fighting in shocked rats served as a coping mechanism, thus preventing the decline in norepinephrine. It would appear that modulation of neurotransmitter levels is stimulus bound, but this relationship is dependent on the availability of coping mechanisms.

The effects of stressors on acetylcholine activity parallel the modifications observed with catecholamines. Stress of electric shock tends to elevate acetylcholine levels (cf. Aprison & Hingtgen, 1970; Aprison, Kariya, Hingtgen, & Toru, 1968; Karczmar et al., 1973). As in the case of norepinephrine, ACh modification is dependent on the availability of coping responses. For example, Karczmar et al. (1973) indicated that stress of an inescapable nature produced increases in ACh of 25% and 8% in the telencephalon and midbrain, respectively. However, when equivalent amounts of shock are received in an escape/avoidance task, a decrease in ACh is observed only in the telencephalon.

The interaction between transmitter modification and availability of coping responses is consistent with the findings regarding the U-shaped avoidance function. That is, the U-shaped function is observed following exposure to inescapable shock or following only a few avoidance training trials. In both of these conditions animals have difficulty in coping with stress, since in the first case there is no availability of learning the response-shock contingencies, and in the second case the learning of these contingencies may require several more trials. As a result, brain norepinephrine decreases and ACh increases. When the stress is of a controllable nature, norepinephrine increases and ACh decreases. The latter set of events lends itself to efficient response initiation and consequently high rates of avoidance performance. In the former case response initiation should be retarded and avoidance should proceed relatively slowly. Moreover, time parameters for the depletion and replenishment of norepinephrine are congruent with the temporal variations of avoidance behavior.

In a recent paper, Broverman, Klaiher, Vogel, and Kobayashi (1974) have for-

warded an argument that in many ways is similar to that presented here. These investigators suggested that in short-term stress situations, release of epinephrine may stimulate central norepinephrine via the reticular activating system and the hypothalamus (see also Gellhorn, 1957), consequently resulting in increased responsivity. However, following prolonged stress, norepinephrine may be replaced by an increase of false transmitters. Accordingly, a state of central cholinergic dominance prevails, with concomitant inhibition of ongoing behaviors. The thesis of the present paper is not in disagreement with this hypothesis, but suggests further that the initial norepinephrine stimulation may result in compensatory increases in ACh. Accordingly, the net result may be one of greater imbalance of the catecholamine-cholinergic systems in favor of cholinergic action. Moreover, if decreased catecholamine-receptor stimulation occurs because of accumulated false transmitters, as suggested by Broverman et al. (1974), a mechanism for increased ACh levels can be postulated, since dopamine apparently has a tonic inhibitory effect on ACh levels (Bartholini et al., 1973).

It is noteworthy that the present formulation does not exclude the involvement of other transmitter systems. Specifically, serotonin (5-HT) has been found to increase following exposure to inescapable shock but not following escapable shock (cf. Karczmar et al., 1973). The fact that 5-HT may be involved in response initiation (Aghajanian & Sheard, 1968; Carlton & Advokat, 1973; Lorens, 1973; Steranka & Barrett, 1974) and in the modulation of ACh and norepinephrine action (Swonger & Rech, 1972) raises the possibility that this neurotransmitter could also contribute to time-dependent changes in avoidance.

Pharmacological Manipulations

A number of studies clearly indicate that alteration in neuronal transmission effectively modifies the time-dependent changes in avoidance behavior. For example, as seen in Figure 2, scopolamine, which functionally blocks ACh action, improves avoidance performance, presumably by reducing

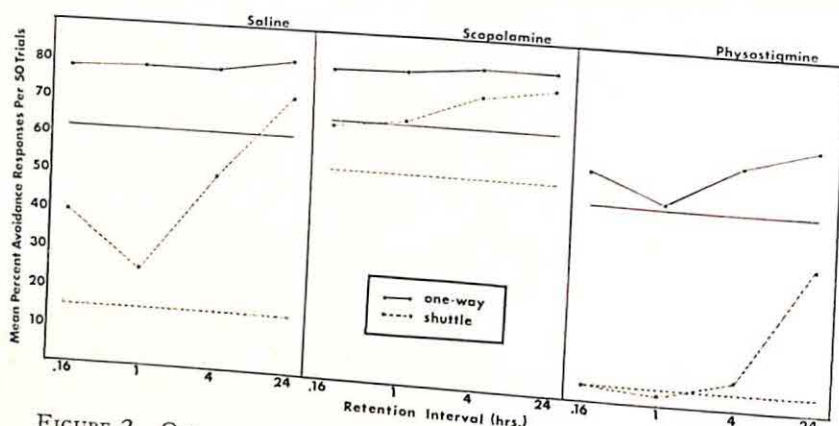


FIGURE 2. One-way and shuttle avoidance performance in Holtzman rats as a function of scopolamine and physostigmine treatment at various intervals following exposure to 10 signaled, inescapable 1 mA shocks. Horizontal solid and dashed lines represent performance of naive rats tested in the one-way and shuttle tasks, respectively. (From "Cholinergic Mechanisms and Alterations in Behavioral Suppression as Factors Producing Time Dependent Changes in Avoidance Performance" by Hymie Anisman, *Journal of Comparative and Physiological Psychology*, 1973, 83, 465-477. Copyright 1973 by the American Psychological Association. Reprinted by permission.)

inhibitory tendencies (Bignami & Rosic, 1971; Carlton, 1969), and also eliminates the U-shaped function. Comparable results have also been observed in several other studies (Anisman & Kokkinidis, 1974; Thompson & Neilson, 1972; Brett, Note 1). Moreover, these investigators have demonstrated that central but not peripheral ACh modification will eliminate the decline in performance at the intermediate interval. Predictably, septal lesions, which reduce cholinergic input to the hippocampus (Lewis, Shute, & Silver, 1967; Lynch, Mathews, Mosko, Parks, & Cotman, 1972; Mellgren & Srebro, 1973), also have been found to effectively reduce the time-dependent variations in avoidance behavior (Anisman & Kokkinidis, 1974). This particular finding might suggest that the septo-hippocampal axis is intricately involved in the U-shaped function; however, direct chemical application to these sites is necessary for further confirmation. It should be noted, however, that application of cholinergic and anticholinergic agents to the septal area does alter avoidance behavior in a predictable manner (see, for example, Grossman, 1964; Kelsey & Grossman, 1969).

In contrast to the effects of anticholinergics, avoidance performance typically has

been noted to decline when animals were treated with drugs (e.g., physostigmine) that increase ACh (Bignami & Gatti, 1966; Bignami et al., 1975). Following treatment with physostigmine, the U-shaped function was eliminated in the shuttle task (see Figure 2, right-hand panel). At the brief and intermediate intervals, the increase in ACh by treatment with physostigmine, together with the shock-induced cholinergic reaction, summate to retard performance. Thus, the differences in performance otherwise seen between the brief and intermediate intervals are precluded. At the long retention interval, the shock-induced ACh changes are minimal, and consequently the disruptive effects of physostigmine on avoidance behavior are diminished. Alternatively, the case could be made that within 24 hours after initial training the memory of the cue-shock contingencies are well consolidated, and as a result performance is less susceptible to disruption by the pharmacologic manipulation.

If we turn to the performance in the one-way avoidance task, it will be remembered that in Holtzman rats a U-shaped function was not observed in the one-way task. It is assumed that ACh levels are highest and response suppression most pronounced at

the intermediate retention interval. However, because freezing behavior can be suppressed readily in the one-way task, differences in behavioral suppression were not manifested in the avoidance task. Administration of physostigmine produced an increase in ACh, which was greater in the 1-hour group than in the remaining groups. Thus, response suppression was increased sufficiently so that several trials were necessary for an effective running response to become established. Under these conditions, differences in response suppression among the various groups were detected in the one-way task. The important point here is that the Treatment $\times$ Retention Interval interaction was not simply due to performance being relatively nonmodifiable at the brief and long intervals due to ceiling effects. The U-shaped function can be eliminated or established by appropriate pharmacological treatments.

With respect to the involvement of the catecholamine system, it has been observed that administration of d-amphetamine improves avoidance behavior (Barrett et al., 1973; Bignami, Robustelli, Janku, & Bovet, 1965; Kreickhaus, Miller, & Zimmerman, 1965) and also eliminates the U-shaped function (Anisman & Kokkinidis, 1974; Barrett & Ray, Note 4). Administration of peripheral adrenergic (Anisman & Kokkinidis, 1974) and cholinergic agents (Anisman & Kokkinidis, 1974; Thompson & Neilson, 1972; Brett, Note 1) do not affect the U-shaped function, clearly indicating that the effects of the catecholamine stimulant and anticholinergic are mediated at a central level. Given the available data, it seems that the catecholamine and cholinergic systems play a major role in determining response suppression, as well as the time-dependent changes in behavior typically observed.

If we turn from the shock-induced suppression to the hyperactivity produced by shock, it appears likely that the latter behavior is mediated by norepinephrine activity (Anisman & Cygan, in press). When animals receive systemic injection of d-amphetamine (1–10 mg/kg in mice), general activity typically is observed to increase.

Following exposure to noncontingent foot shock the extent of the hyperactivity is greatly magnified. The excitation of behavior appears to involve central catecholamine activity, since the peripherally acting agent, p-hydroxyamphetamine, does not produce postshock excitation. When animals are pretreated with an inhibitor of dopamine-B-hydroxylase, bis-(4-methyl-1-homopiperazinylthiocarbonyl) disulfide, the enzyme responsible for the conversion of dopamine to norepinephrine, general activity and amphetamine-induced locomotor activity are not affected. Apparently, the locomotor stimulant properties of d-amphetamine are due to dopamine rather than norepinephrine activity (see, for example, Carlsson, 1970; Fuxe & Ungerstedt, 1970). However, pretreatment with the dopamine-B-hydroxylase inhibitor entirely eliminated the amphetamine-shock synergism. Together with earlier reports indicating that shock potentiates the turnover of norepinephrine induced by d-amphetamine (Javoy et al., 1968), these data suggest that changes in norepinephrine turnover subserve the behavioral excitation produced by shock.

It is noteworthy that the postshock excitation was not apparent with treatment of scopolamine, although the anticholinergic effectively reduced the suppression of activity produced by shock. Apparently, the initial rise of activity is probably specific to effects on catecholamine activity. Thus, although balance between the catecholaminergic and cholinergic systems may exist, this balance may be one in which the cholinergic system mediates response inhibition/disinhibition, whereas the effects of catecholamine activity are in excess of simple response disinhibition.

Associative Versus Nonassociative Effects of Neurotransmitter Modification

It might be argued from the standpoint of a memory hypothesis that although neurotransmitters are involved in the production of the U-shaped function, this involvement is not at the nonassociative level, but rather directly affects memory processes. However, a memory interpretation for the U-

shaped function is supported neither by the data from a purely behavioral vantage nor by studies evaluating the source of the pharmacological effects on avoidance behavior. To be more explicit, it has been posited (Klein, 1972; Klein & Spear, 1970a, 1970b, 1973; Spear, 1973; Spear, Klein, & Riley, 1971) that the U-shaped function is due to retrieval failure at an intermediate interval, owing to time-dependent variations in endogenous neurotransmitters and steroids. Essentially, it was suggested that stress results in numerous chemical changes (e.g., adrenocorticotrophic hormone, corticosterone, epinephrine, norepinephrine, acetylcholine, as well as other unnamed endogenous substances). These changes produce state-dependent (drug dissociated) effects, such that the organism cannot retrieve the original memories unless the organism's state during testing is comparable to that which existed during original training. It is supposed that soon after shock the endogenous chemical changes normally elicited by a stress experience are not yet drastic, and consequently memory remains relatively stable. One hour after the stress experience, the endogenous chemical changes are maximal, and as a result memory retrieval is impaired. Within 24 hours, the internal state of the animal returns to a basal level and the organism again can retrieve adequately the previous learned response.

The major source of support for the memory hypothesis has been based on several experiments conducted by Klein and Spear (1970a, 1970b), which indicated that when animals were trained on an active avoidance task and tested on a passive avoidance task (or vice versa), negative transfer was observed at a brief or a long interval relative to naive animals. However, at an intermediate interval, performance was comparable to that of naive animals. The conclusion reached by these investigators is that at the intermediate interval the organism is unable to retrieve the previously established response-shock contingency, precluding the possibility of negative transfer arising. When the previously established memory is intact, as at the brief or long interval, negative transfer arises because of the con-

flicting response requirements of the two tasks.

Paradoxically, employing a somewhat similar type of test situation, several investigators (e.g., Anisman & Waller, 1971; Barrett, Leith, & Ray, 1971b; Bauer, 1973; de Toledo & Black, 1970; Steranka & Barrett, 1973) have observed substantially different results. These investigators shocked animals in a compartment of a particular color, and subsequently had independent groups of rats run either toward or away from the preshock compartment at one of various intervals. In both the conditions in which subjects were required to run towards or away from the preshock stimuli, a U-shaped performance function was observed. Not surprisingly, animals required to run away from those cues previously associated with shock acquired the avoidance response more readily than animals required to run toward the preshock stimuli. Of prime importance here is the finding that this difference was observed regardless of the test-retest interval. Moreover, the data presented by de Toledo and Black (1970) and Steranka and Barrett (1973) also indicated that performance at the intermediate interval was inferior to that observed in naive animals. If memory could not be retrieved at the intermediate interval, as suggested by Klein and Spear (1970a, 1970b), then it should not have mattered whether animals were required to run away or toward the preshock stimuli, and consequently both groups should have performed equally poorly at the intermediate interval. In addition, performance at the intermediate interval should not have differed from that of naive animals.

In this same vein, Anisman (1973a) has observed that if relatively intense shock (2 mA) was employed during initial training, then performance at an intermediate interval was poorer than that of a naive baseline group. Had memory been disrupted, then performance at the intermediate interval should have equalled, or perhaps exceeded, that observed in naive animals. Together with the data reported by Barrett et al. (1971b) and Steranka and Barrett (1973), in which avoidance, but not discrimination

performance, was found to vary over retention intervals, these results suggest that response initiation abilities rather than retrieval failure are responsible for the avoidance deficit seen at the intermediate interval. This is not to say that associative factors are entirely unimportant in the production of the U-shaped function, but rather that changes in response initiation appear to account for a greater portion of the variance than do potential changes in retrieval abilities.

It has been suggested, concerning the involvement of steroids, that since corticosterone levels vary following the inception of a stressor and are differentially affected by the extent of control animals have over a stressor (e.g., Weiss, 1971a, 1971b, 1971c), this steroid might mediate the time-dependent changes in performance (Brush & Levine, 1966; Levine & Brush, 1967; Singh, Sakallaris, & Brush, 1971). In support of this notion Brush and Levine (1966) reported that over time following fear conditioning, alterations in corticosterone correlated with avoidance behavior. The depletion of corticosterone is thought to disrupt avoidance, either because animals are unable to cope with a subsequent stress or because of failure in memory retrieval (Brush, 1971). However, contrary to such a position, several studies have failed to find any effect of pituitary-adrenal manipulations on the U-shaped function (Barrett et al., 1971b; Kasper-Pandi, Hansing, & Usher, 1970; Snider, Marquis, Black, & Suboski, 1971; Suboski, Marquis, Black, & Platenius, 1970), suggesting that the relationship between corticosterone level and the decline in avoidance performance at an intermediate interval is not a causal one. It is quite likely that hypothalamic norepinephrine plays a tonic inhibitory role on the corticotropin-releasing factor. Thus, with the decline in norepinephrine following the introduction of a stressor, corticosterone is released (Scapagnini, Annunziato, & Preziiosi, 1973). Predictably, the corticosterone release, like that of norepinephrine, would be dependent on whether the stress is one that the animal can cope with. In any event, the important point for the present

purpose is that the U-shaped function is not causally linked to changes in adrenal steroids.

If we turn to the effects of drug treatments on the modification of behavior, it seems unlikely that the pharmacological treatments that alter the U-shaped function affect memory processes. For example, Barrett et al. (1973) observed that although scopolamine and d-amphetamine improved avoidance performance in a Y-maze, there were no changes in discrimination learning as a result of the drug treatment. Evidently, the pharmacological treatment affects response-modulating (i.e., disinhibitory) systems as opposed to memory processes. Employing a somewhat different approach, Bignami and his associates (Bignami, Amorico, Frontali, & Rosic, 1971; Bignami & Rosic, 1971; Rosic & Bignami, 1970) reported data that have lent substantial credence to a response modulation hypothesis. These investigators observed that in a go-no-go situation, scopolamine increased errors of commission but decreased errors of omission. The drug treatment apparently decreased response inhibition, thereby increasing the probability of active responses being made and concomitantly diminishing the effectiveness of one component of the passive avoidance response. Results recently obtained in our laboratory (Anisman, *in press*) have replicated Barrett's earlier findings that scopolamine and d-amphetamine improved active Y-maze avoidance but had negligible effects on discrimination performance. Predictably, both agents tended to disrupt inhibitory avoidance behavior. Had the drug treatments altered memory, then facilitative effects should have been observed in active avoidance, inhibitory avoidance, and discrimination performance. Together, these experiments, as well as additional data reviewed by Bignami (*in press*) and Carlton and Markiewicz (1971), indicate that the actions of scopolamine and d-amphetamine are task specific in that they affect inhibitory tendencies.

The notion that scopolamine and d-amphetamine selectively modulate tasks involving inhibitory tendencies has been verified in a series of experiments designed to examine just such a possibility. Specifically, Suits

and Isaacson (1968; see also Anisman, 1973a) reported that the facilitative effects of the drug treatments were apparent in a shuttle task but not in a one-way task. In the shuttle task, where freezing is relatively high, scopolamine improves performance through the reduction of freezing behavior. In contrast, when freezing ordinarily can be suppressed readily, as is the case in the one-way task, beneficial effects elicited by further elimination of response suppression are negligible. In a similar vein, Oliverio (1967) has reported that the beneficial effects of scopolamine are evident early in training, when freezing presumably is relatively high, but not later in training, when freezing has been suppressed to some extent. Finally, data reported by Anisman (1975; for complete reviews of relevant literature, see Bignami, in press; Bignami & Rosic, 1971; Carlton, 1969; Carlton & Markiewicz, 1971; Longo, 1966) also indicated that if the task is very difficult, the beneficial effects of scopolamine and amphetamine may not be apparent. It seems that the presence of a strong disruptive associative component may override the response-modulating effects of the drug treatments. In effect, the mere increase in response-initiating abilities does not guarantee success in an avoidance task if the signal-shock and response-shock contingencies are not learned quickly. If the task is one in which the response and shock contingencies are easily acquired and avoidance is limited by response inhibition, then the pharmacological treatments may improve performance. However, the drugs do not appear to alter associative processes *per se*.

While the basic premise of the present paper is that nonassociative processes are mediated by neurochemical events, it is necessary to interject several caveats. In particular, the argument may be advanced that although noncontingent shock is followed by behavioral and neurochemical changes, the relationship between these events is simply correlational and does not imply causality. The stress situation could set in motion unknown mechanisms that result in response repertoire modification and quite independently produce neurochemical changes. Alternatively, the re-

sponse repertoire changes themselves may elicit neurochemical modification. The fact that pharmacological manipulations result in predictable changes in behavior argues against such positions and suggests a causal relationship between neurochemical activity and nonassociative processes. Nevertheless, at this time it is premature to dismiss entirely the possibility that associative processes or response-repertoire changes influence neurochemical events.

A second caveat, discussed earlier with respect to time-dependent changes in performance, needs to be reiterated here in terms of the effects of pharmacological treatments. Specifically, although motor enhancement or response disinhibition favor improved avoidance performance, pharmacological treatments that eliminate these non-associative factors do not insure high levels of avoidance. In the presence of adequate associative processes, treatments such as scopolamine will enhance performance if response suppression is the limiting factor for high avoidance rate. In some tasks the response enhancing or disinhibitory qualities of a drug may disrupt rather than augment performance. For example, several reports are available indicating that scopolamine may disrupt one-way and pole-climb avoidance (cf. Bignami et al., 1965). These contradictions apparently are a consequence of the response-repertoire change induced by the drug, together with the response requisite of the avoidance task. In the one-way task, incompatible responses established by shock are suppressed quickly. Thus, an increase in exploration produced by scopolamine may compete with effective defensive behaviors that occur otherwise. Similarly, in the case of pole climbing, the increase of motor activity may well be incompatible with the climbing/grasping behavior necessary for efficient avoidance. It seems that in order to analyze the behavioral effects of drugs, it is not sufficient simply to view the response biases created by the pharmacological treatments, but rather the conflicts between these biases together with the various components of the avoidance task should be evaluated. Indeed, precisely such formulations yield additional important evidence regarding the

role of coping behavior and response-repertoire changes in the modulation of behavior following drug treatment. In the case of passive avoidance, coping behavior is possible, since specific response-shock contingencies exist. Although acetylcholine levels should not increase under these conditions (see earlier discussion), scopolamine effectively eliminates inhibitory tendencies and disrupts performance (cf. Bignami & Rosic, 1971). Clearly, the effects of the drugs are dependent on response styles rather than on correlates of coping behavior per se.

Finally, one additional limiting condition to the role of neurotransmitter modification and its relationship to performance should be introduced. When considering the interactions among coping behavior, neurotransmitter changes, and response requisites, it is necessary to recognize that any predictions of behavior are dependent on species and strain differences. In addition to differences in shock-induced suppression, various strains of mice have been found to differ in response to drug treatments and the interaction between drug treatments and shock (Anisman & Cygan, in press; Anisman et al., in press). Similarly, strains of mice have been found to exhibit varied levels of catecholamines and acetylcholine as well as enzymes associated with these transmitters (Ebel, Hermetet, & Mandel, 1973; Pryor, Schlesinger, & Calhoun, 1966). Given the existing strain variations in neurotransmitter activity and in response to stress, it should not be surprising to observe strain specific temporal variations of the nonassociative effects of shock (Anisman, 1973a). Similarly, since different strains of rats have been found to exhibit different response styles (in terms of anticipatory responding) following noncontingent inescapable shock, it is predictable that strains exhibiting less anticipatory responding would exhibit poorer active avoidance behavior (Katzev & Mills, 1974) and greater disruption of performance after prolonged exposure to signaled inescapable/unavoidable shock (Shurman & Katzev, 1975).

In summary, strain differences may contribute greatly to the effects of shock on performance via inherent differences in terms of their response to aversive stimula-

tion. At the same time a genetic/biochemical approach in the evaluation of behavior phenomena may be useful in drawing fracture lines between the roles of response contingencies, coping style, and neurotransmitter activity in the modulation of performance. That is, it should be possible to evaluate the effects of a particular treatment on strains exhibiting similar response styles in the face of contingent or noncontingent stress but differing in neurochemical activity. Under these conditions, or permutations of these conditions, it may be possible to evaluate the role of a variety of factors, at both an associative and nonassociative level, in determining ultimate response rates (cf. Anisman, 1975, in press; Anisman & Cygan, in press; Anisman et al., in press). While at present there appears to be a central core of correlates of nonassociative factors, such as variations in neurotransmitter activity and stress-induced excitation and inhibition of behavior, this definition must be considered provisional, pending evaluation of other potential concomitants of contingent stress, for example, coping by anticipatory responding (Katzev & Mills, 1974) and pseudoconditioning (Izquierdo, 1974).

Extensions and Implications

The model outlined in the previous sections is applicable to other stress-related phenomena. Some of these are presented below; additional pertinent implications are available in Gellhorn (1957).

One particularly fascinating phenomenon that has received considerable attention over the past few years is the finding that animals subjected to inescapable/unavoidable aversive stimulation subsequently exhibit impoverished escape/avoidance performance (cf. Anisman & Waller, 1972, 1973; Maier, Seligman, & Solomon, 1969; Seligman, Maier, & Solomon, 1971). Seligman and his associates have attributed the disruptive influence of the initial stress to "learned helplessness." Specifically, in the face of inescapable unavoidable stress, animals learn that their responses are independent of shock offset. Accordingly, in subsequent encounters with similar stressors, the organism neither attempts to escape nor to avoid the

aversive stimuli. In support of this hypothesis it was observed that if animals were first trained in an avoidance task and then subjected to the inescapable shock procedure, subsequent avoidance performance was not disrupted. In effect, the suggestion by these investigators is that if animals learn that control is possible, helplessness will not result from the inescapable shock procedure.

As an alternative to such a model, several investigators have suggested that the disruptive effects of inescapable shock are due to either associative (Weiss, Kreickhaus, & Conte, 1968) or nonassociative (Anisman & Waller, 1971, 1972, 1973b) competing responses elicited by the initial stress experience. Moreover, Weiss et al. (1970) have suggested that the freezing response may be mediated by decreased levels of norepinephrine, owing to the severity of the shock procedure used, together with the lack of coping responses available to the animal (see previous discussion). When animals have some control over the stressor, then norepinephrine levels are not greatly modified, and consequently disruption in performance is not observed. Clearly, the two hypotheses are quite similar; the latter assuming that ultimately, nonassociative effects mediated by neurochemical changes are responsible for the disruptive effects of inescapable shock, while the helplessness hypothesis attributes the inescapable shock effects to cognitive factors without discarding biochemical models.

Several specific predictions follow from the neurochemical model:

1. Pharmacological treatments that increase norepinephrine levels should mitigate the disruptive influence of inescapable shock, whereas the converse should be true of agents that deplete norepinephrine.
2. If the cholinergic system is involved in the mediation of the effects of inescapable shock, either via a rebound from adrenergic action or through direct effects of shock, then modification of ACh levels should promote predictable changes in behavior following inescapable shock.
3. If the disruptive effects of the inescapable shock procedure are mediated primarily by neurochemical modification and not associative processes per se (i.e., learning that

shock offset is independent of responses), then it follows that disruption in performance should be modifiable through pharmacologic treatments, without resort to the inescapable shock procedure. Moreover, if disruption in performance under such conditions is observed, then these effects ought to be subject to modification in the same way as the inescapable shock effects are modifiable. For example, prior avoidance training should immunize the animal against drug-induced disruption of performance.

Much of the currently available data are in accord with these predictions. For example, it was observed that d-amphetamine (Anisman & Kokkinidis, 1974) or scopolamine (Anisman, 1973a) eliminated the disruptive effects of inescapable shock, whereas increasing ACh levels through treatment with physostigmine exacerbated the disruptive effects of the inescapable shock procedure (Anisman, 1973a). Similarly, Thomas and Balter (in press) found that direct application of an anticholinergic to the septal area alleviated the disruption of avoidance induced by the initial stress procedure. Clearly, the effects of inescapable shock are modifiable via pharmacological manipulations. Of prime importance, however, is that both central catecholamine or cholinergic manipulations produce similar effects, thus implicating both of these systems in the involvement of the disruptive influences of inescapable shock. Inasmuch as the pharmacological agents employed do not have profound influences on associative processes (*vis-à-vis* earlier discussion), it is likely that nonassociative factors mediate the poor performance following the inescapable shock procedure.

With respect to the effects of drug treatments that increase ACh or decrease catecholamine levels in the absence of a shock procedure, the available data must be treated with caution, since with higher dosages pronounced disruption of motoric abilities are produced. Accordingly, only those studies in which locomotor disruption is not apparent are applicable to the present analysis. Moreover, in order that a drug treatment be pertinent to the current discussion, the drug must be capable of mimicking the ef-

fects typically observed in the inescapable shock paradigm. With these stipulations borne in mind, there are several reports available that suggest that the effects of inescapable shock may be mimicked simply via pharmacological treatments. For example, Weiss et al. (in press) reported that reduction of catecholamine levels with tetra-benzazine produced behavior which resembled closely that of inescapable shock treatments, in that disruption was observed in a fairly difficult shuttle task but not in a simpler task requiring a nose-pressing response. Similarly, Fibiger, Phillips, and Zis (1974) observed that 6-hydroxydopamine (6-OHDA) lesions of the substantia nigra, which reduced norepinephrine and dopamine levels, also retarded the acquisition of an active avoidance response. However, if the response had been established previously, no disruption in performance was produced by the 6-OHDA treatment.

Consistent with the data regarding decreased levels of catecholamines, it has been observed consistently that increasing ACh levels results in disruption of performance (cf. Bignami & Gatti, 1966; Bignami et al., 1975). As in the inescapable shock paradigm, the disruption of shuttle avoidance with increased levels of ACh was substantially greater than that seen in a one-way avoidance task (Anisman, 1973a). Furthermore, it seems to be the case that the effects of agents that increase ACh levels are more potent in animals that show low baseline levels of performance than in animals with higher and more stable baselines (Anisman, 1973a; Rosic, Bignami, & Gatti, 1971), and less potent on already established avoidance responses than on novel avoidance responses (cf. Anisman, 1973a; Bignami et al., 1975).

Taken together, it seems that the effects of inescapable shock on avoidance can be mimicked by treatments that decrease catecholamine levels or increase ACh levels, suggesting that the disruptive effects of inescapable shock may be mediated by variation in neurotransmitter levels. This is not to say that "helplessness," as suggested by Seligman, is not a viable position. Clearly, neurotransmitter alterations are dependent

on whether or not animals can cope with the stressor. Rather, the suggestion here is that the invocation of a cognitive model, at least in the animal subject, is superfluous to the understanding of the disruptive effects of inescapable shock.

Recently, Seligman and his associates (Miller & Seligman, 1973; Seligman, 1974; Seligman, Klein, & Miller, in press) suggested that the helplessness notion might serve as a useful analogue for reactive depression in humans, in that many of the symptomatologies appear similar. The hazards inherent in such generalizations are well known. Nevertheless, the suggestion is a tantalizing one. Inasmuch as depression is typically modified with treatments that increase catecholamine activity (cf. Schildkraut, 1965), as well as those treatments that decrease cholinergic activity (Janowsky et al., 1972; see also Lloyd & Hornykiewicz, 1974, for review of this literature), it is likely that these neurotransmitters are involved in the mediation of depressive behaviors. Whether or not the biochemical changes are in fact due to uncontrollable stresses is not clear. However, if a model of depression involving the helplessness notion is to be employed, it would be advantageous to include the neurotransmitter notion, and especially that of balanced transmitter systems, into these formulations. Such a position also is amenable to other currently existing models, such as those which suggest that genetic influences may be involved in predisposing individuals to depression (Akiskal & McKinney, 1973; Winokur, 1973). Thus, one might speculate that inherited biochemical levels is the underlying phenotype regarding the behavioral response to various stressors. Again, a cautionary note needs to be interjected here, namely, although pharmacological intervention produces beneficial results in depressive behaviors, suggesting a causal link between chemical levels and the behavior, the possibility still exists that the neurotransmitter activity may come under cognitive control or is affected by cognitive state. Moreover, in dealing with stress/biochemical relations it is necessary to define what represents inadequate coping in differ-

ent individuals, and other things being equal, how do different individuals react, in terms of biochemical changes, to a standard difficulty in coping behavior.

Finally, another stress-related occurrence in which rebound effects may play a major role is that of gastric ulceration. Specifically, Weiss (1971a, 1971b, 1971c) has observed that in those stress situations in which feedback is unavailable, animals do not have methods for coping with stress, and when shock is made contingent on appropriate coping responses, large amounts of ulceration typically are observed. Each of these situations has been noted to be associated with low levels of norepinephrine. Inasmuch as ulceration typically occurs only after several hours following stress termination (Desiderato, Mackinnon, & Hissom, 1974), it is likely that the initial stress-induced norepinephrine reaction is countered by both central and peripheral cholinergic action. The parasympathetic rebound would be expected to result in vagal innervation and hydrochloric acid secretion, thereby promoting ulceration. This notion is clearly speculative, however, and requires further evaluation.

While the present paper has suggested that rebound effects may be responsible for several stress-related phenomena, it should be clear that such a notion may be amenable to other clinical phenomena. For example, Gellhorn (1957) has suggested that phenomena such as functional psychosis, emotional responsivity, and hormonal regulation might involve a rebound type of model. Similarly, it is likely that other paradoxical and poorly understood phenomena, such as the enigmatic effects of sympathomimetics in hyperkinetic children or dual effects of lithium in manic and depressive states, involve the innervation of antagonistic systems following excessive stimulation.

Concluding Remarks and Summary

While associative factors clearly act as a rate-limiting step in the acquisition of aversively motivated behaviors, the suggestion of the present paper is that nonassociative factors similarly determine avoidance response rate. Needless to say, in the absence of

associative abilities, reduction of incompatible nonassociative factors will not enhance learning. However, in the presence of good associative abilities, nonassociative effects elicited by a stressor may hinder performance. When these nonassociative effects are eliminated, the observed response rate is a truer indication of learned behaviors.

Within aversively motivated paradigms, nonassociative factors are largely mediated by variation in neurotransmitter activity. Reduction in norepinephrine and/or dopamine or increases in acetylcholine result in increased response inhibition, which augments passive avoidance but disrupts active responding. Modification of transmitter levels should effectively alter performance level if poor performance is a result of non-associative factors. Although behaviors may be altered via manipulations of a particular transmitter, it is likely that several transmitters act jointly to determine the ultimate response rate. Basically, the suggestion was made that catecholamines and acetylcholine act in a balanced antagonistic manner, thereby maintaining neurochemical homeostasis. However, when the balance moves in favor of cholinergic dominance, active responding is inhibited in aversively motivated paradigms.

In addition to the existence of balanced transmitter systems, it is likely that with excessive stimulation of one system, an antagonistic rebound may occur in a second system in order to maintain neurochemical stability. Paradoxical behavioral effects may occur, however, since the rebound may be a sluggish one in terms of onset and termination, while the initial transmitter change elicited by direct stimulation may be rapid in onset and transient in duration. The inclusion of such a model adequately explains several time-dependent phenomena associated with aversive stimulation, as well as phenomena dealing with traumatic stresses and factors dealing with control over stress. Finally, it can be applied as a model to clinically related behaviors.

REFERENCE NOTES

1. Brett, C. W. *Acquisition and retention of an active avoidance response to intense white noise: Kamin effect*. Unpublished master's thesis, 1973.

- (Information concerning this thesis is available from H. Anisman, Department of Psychology, Carleton University, Ottawa, Canada.)
2. Anisman, H., & Kokkinidis, L. *Reversal of d-amphetamine induced locomotor retardation by scopolamine*. Unpublished manuscript, 1975. (Available from H. Anisman, Department of Psychology, Carleton University, Ottawa, Canada.)
 3. Weiss, J. M., Pohorecky, L. A., Dorros, K., & Williams, S., Emmel, D., Whittlesey, M., & Case, E. *Coping behavior and brain norepinephrine turnover*. Paper presented at the meeting of the Eastern Psychological Association, Washington, D. C., May 1973.
 4. Barrett, R. J., & Ray, O. S. *The effect of adrenergic-cholinergic manipulations on time dependent processes in avoidance learning*. Paper presented at the meeting of the Psychonomic Society, St. Louis, November 1971.
- ### REFERENCES
- Aghajanian, G. K., & Sheard, M. H. Behavioral effects of midbrain raphe stimulation. *Communications in Behavioral Biology*, 1968, 1, 31-41.
- Agid, Y., Guyenet, P., Javoy, F., Beaujouan, J. C., & Glowinski, J. Specific aspects of antagonists and agonists of DA receptors on ACh turnover in the neostriatum. *Journal of Pharmacology*, 1974, 5, 59.
- Ahtee, L., & Kääriäinen, I. The role of dopamine in pilocarpine-induced catelepsy. *Naumynine in pilocarpine-induced catelepsy*. *Schmiedeberg's Archives of Pharmacology*, 1974, 289, 25-38.
- Akiskal, H. S., & McKinney, W. T. Depressive disorders: Toward a unified hypothesis. *Science*, 1973, 182, 20-29.
- Anden, N. E. Dopamine turnover in the corpus striatum and the limbic system after treatment with neuroleptic and anti-acetylcholine drugs. *Journal of Pharmacy and Pharmacology*, 1972, 24, 905-906.
- Anden, N. E., & Bedard, P. Influences of cholinergic mechanisms on the function and turnover of brain dopamine. *Journal of Pharmacy and Pharmacology*, 1971, 23, 460-462.
- Anisman, H. Cholinergic mechanisms and alterations in behavioral suppression as factors producing time dependent changes in avoidance performance. *Journal of Comparative and Physiological Psychology*, 1973, 83, 465-477. (a)
- Anisman, H. Effects of pretraining compatible and incompatible responses on subsequent one-way and shuttle-avoidance performance in rats. *Journal of Comparative and Physiological Psychology*, 1973, 82, 95-104. (b)
- Anisman, H. Acquisition and reversal learning of an active avoidance response in three strains of mice. *Behavioral Biology*, 1975, 14, 51-58.
- Anisman, H. Differential effects of scopolamine and d-amphetamine on avoidance behaviors in three strains of mice. *Pharmacology Biochemistry and Behavior*, in press.
- Anisman, H., & Cygan, D. Central effects of scopolamine and d-amphetamine on locomotor activity: Interaction with strain and stress variables. *Neuropharmacology*, in press.
- Anisman, H., & Kokkinidis, L. Effects of central and peripheral adrenergic and cholinergic modification on time dependent processes in avoidance performance. *Behavioral Biology*, 1974, 10, 161-171.
- Anisman, H., & Wahlsten, D. Response initiation and directionality as factors influencing avoidance performance. *Journal of Comparative and Physiological Psychology*, 1974, 87, 1119-1128.
- Anisman, H., Wahlsten, D., & Kokkinidis, L. Effects of d-amphetamine and scopolamine on activity before and after shock in three mouse strains. *Pharmacology Biochemistry and Behavior*, in press.
- Anisman, H., & Waller, T. G. Effects of conflicting response requirements and shock compartment confinement on the Kamin effect in rats. *Journal of Comparative and Physiological Psychology*, 1971, 77, 240-244.
- Anisman, H., & Waller, T. G. Facilitative and disruptive effects of prior exposure to shock on subsequent avoidance performance. *Journal of Comparative and Physiological Psychology*, 1972, 78, 113-122.
- Anisman, H., & Waller, T. G. Effects of inescapable shock on subsequent avoidance performance: Role of response repertoire changes. *Behavioral Biology*, 1973, 9, 331-355. (a)
- Anisman, H., & Waller, T. G. Footshock produced excitation and inhibition of activity in rats. *Animal Learning and Behavior*, 1973, 1, 93-95. (b)
- Aprison, M. H., & Hingtgen, J. N. Evidence of a central cholinergic mechanism functioning during drug induced excitation in avoidance behavior. In E. Heilbron & A. Winter (Eds.), *Drugs and cholinergic mechanisms in the CNS*. Stockholm: Forsvarets Forskningsanstalt, 1970.
- Aprison, M., Kariya, T., Hingtgen, J., & Toru, M. Changes in acetylcholine, norepinephrine and 5-hydroxytryptamine concentrations in several discrete brain areas of the rat during behavioral excitation. *Neurochemistry*, 1968, 15, 1131-1139.
- Arnfred, T., & Randrup, A. Cholinergic mechanism in brain inhibiting amphetamine-induced stereotyped behavior. *Acta Pharmacologica et Toxicologica*, 1968, 26, 384-394.
- Barchas, J. D., Ciaranello, R. D., Stolk, J. M., Keith, H., Brodie, H., & Hamburg, D. A. Biogenic amines and behavior. In S. Levine (Ed.), *Hormones and behavior*. New York: Academic Press, 1972.
- Barchas, J. D., & Freedman, D. X. Brain amines: Response to physiological stress. *Biochemical Pharmacology*, 1963, 12, 1232-1235.
- Barrett, R. J., Leith, N. J., & Ray, O. S. The effects of pituitary-adrenal manipulations on time-dependent processes in avoidance learning. *Physiology & Behavior*, 1971, 7, 663-665. (a)

- Barrett, R. J., Leith, N. J., & Ray, O. S. Kamin effect in rats: Index of memory or shock-induced inhibition. *Journal of Comparative and Physiological Psychology*, 1971, 77, 234-239. (b)
- Barrett, R. J., Leith, N. J., & Ray, O. S. Permanent facilitation of avoidance behavior by d-amphetamine and scopolamine. *Psychopharmacologia*, 1972, 25, 321-331.
- Barrett, R. J., Leith, N. J., & Ray, O. S. A behavioral and pharmacological analysis of the variables mediating active avoidance behavior in rats. *Journal of Comparative and Physiological Psychology*, 1973, 82, 489-500.
- Bartholini, G., Stadler, H., & Lloyd, K. G. Cholinergic-dopaminergic relation in different brain structures. In E. Usdin & S. H. Snyder (Eds.), *Frontiers in catecholamine research*. Oxford, England: Pergamon Press, 1973.
- Bauer, R. H. Brain norepinephrine and 5-hydroxytryptamine as a function of time after avoidance training and footshock. *Pharmacology Biochemistry and Behavior*, 1973, 1, 615-618.
- Benton, D., Moffat, I. I., Rick, J. T., & Taberner, P. V. The influence of 6-hydroxydopamine on mouse brain acetylcholinesterase and glutamic acid decarboxylase activity. *Biochemical Pharmacology*, 1975, 24, 426-428.
- Bignami, G. Nonassociative explanations of behavioral changes induced by central cholinergic drugs. *Acta Neurobiologiae Experimentalis*, in press.
- Bignami, G., Amorico, L., Frontali, M., & Rosic, N. Central cholinergic blockade and two-way avoidance acquisition: The role of response disinhibition. *Physiology & Behavior*, 1971, 7, 461-470.
- Bignami, G., & Gatti, G. L. Neurotoxicity of anticholinesterase agents. Antagonistic action of various centrally acting drugs. *Excerpta Medica International Congress*, 1966, 118, 93-106.
- Bignami, G., Robustelli, F., Janku, I., & Bovet, D. Action de l'amphetamine et de quelques agents psychotropes sur l'acquisition d'un conditionnement de fuite et d'évitement chez des rats sélectionnés en fonction du niveau particulièrement bas de leurs performances. *Compte Rendu de l'Académie des Sciences, Paris*, 1965, 260, 4273-4278.
- Bignami, G., & Rosic, N. The nature of disinhibitory phenomena caused by central cholinergic (muscarinic) blockage. In O. Vinar, Z. Voltava, & P. B. Bradley (Eds.), *Advances in neuropharmacology*. Amsterdam: North-Holland, 1971.
- Bignami, G., Rosic, N., Michalek, M., Milosevic, S., & Gatti, G. L. Behavioral toxicity of anticholinesterase agents: Methodological, neurochemical, and neuropsychological aspects. In B. Weiss & V. G. Laties (Eds.), *Behavioral toxicology*. New York: Plenum Press, 1975.
- Bliss, E. L., Ailion, J., & Zwanziger, J. Metabolism of norepinephrine, serotonin, and dopamine in rat brain with stress. *Journal of Pharmacology and Experimental Therapeutics*, 1968, 164, 122-134.
- Bliss, E. L., & Zwanziger, J. Brain amines and emotional stress. *Journal of Psychiatric Research*, 1966, 4, 189-198.
- Bolles, R. C. Species-specific defense reactions and avoidance learning. *Psychological Review*, 1970, 77, 32-48.
- Bolles, R. C. Species-specific defense reactions. In F. R. Brush (Ed.), *Aversive conditioning and learning*. New York: Academic Press, 1971.
- Bowers, M. B., & Roth, R. H. Interaction of atropine-like drugs with dopamine-containing neurones in rat brain. *British Journal of Pharmacology*, 1972, 44, 301-306.
- Breese, G. R., Cooper, B. R., & Smith, R. D. Biochemical and behavioral alterations following 6-hydroxydopamine administration into brain. In E. Usdin & S. H. Snyder (Eds.), *Frontiers in catecholamine research*. Oxford, England: Pergamon Press, 1973.
- Breitner, C., Picchioni, A., & Chin, L. Neurohormone levels in brain after CNS stimulation including electrotherapy. *Journal of Neuropsychiatry*, 1963, 5, 153.
- Broverman, D. M., Klaiber, E. L., Vogel, W., & Kobayashi, Y. Short-term versus long-term effects of adrenal hormones on behaviors. *Psychological Bulletin*, 1974, 81, 672-694.
- Brush, F. R. Avoidance learning as a function of time after fear conditioning and unsignalled shock. *Psychonomic Science*, 1966, 1, 405-406.
- Brush, F. R. Retention of aversively motivated behavior. In F. R. Brush (Ed.), *Aversive conditioning and learning*. New York: Academic Press, 1971.
- Brush, F. R., & Levine, S. Adrenocortical activity and avoidance learning as a function of time after fear conditioning. *Physiology & Behavior*, 1966, 1, 309-311.
- Butcher, L. L., & Engel, J. Peripheral factors in the mediation of the effects of L-dopa on locomotor activity. *Journal of Pharmacy and Pharmacology*, 1969, 21, 614-616.
- Campbell, B. A., Lytle, L. D., & Fibiger, H. C. Ontogeny of adrenergic arousal and cholinergic inhibitory mechanisms in the rat. *Science*, 1969, 166, 635-637.
- Campbell, B. A., & Mabry, P. D. The role of catecholamines in behavioral arousal during ontogenesis. *Psychopharmacologia*, 1973, 31, 253-264.
- Carlsson, A. Drugs which block the storage of 5-hydroxytryptamine and related amines. In V. Erspamer (Ed.), *5-Hydroxytryptamine and related indolealkylamines (Handbook of experimental pharmacology, Vol. 19)*. Berlin: Springer-Verlag, 1966.
- Carlsson, A. Amphetamine and brain catecholamines. In E. Costa & S. Garattini (Eds.), *Amphetamines and related compounds*. New York: Raven Press, 1970.
- Carlton, P. L. Augmentation of the behavioral effects of amphetamine by atropine. *Journal of*

- Pharmacological and Experimental Therapeutics*, 1961, 132, 91-96. (a)
- Carlton, P. L. Augmentation of the behavioral effects of amphetamine by scopolamine. *Psychopharmacologia*, 1961, 2, 377-380. (b)
- Carlton, P. L. Cholinergic mechanisms in the control of behavior by the brain. *Psychological Review*, 1963, 70, 19-39.
- Carlton, P. L. Brain acetylcholine and habituation. *Progress in Brain Research*, 1968, 28, 48-60.
- Carlton, P. L. Brain-acetylcholine and inhibition. In J. T. Tapp (Ed.), *Reinforcement and behavior*. New York: Academic Press, 1969.
- Carlton, P. L., & Advokat, C. Attenuated habituation due to parachlorophenylalanine. *Pharmacology Biochemistry and Behavior*, 1973, 1, 657-663.
- Carlton, P. L., & Didamo, P. Augmentation of the behavioral effects of amphetamine by atropine. *Journal of Pharmacological and Experimental Therapeutics*, 1961, 134, 91-96.
- Carlton, P. L., & Markiewicz, C. Behavioral effects of atropine and scopolamine. In E. Furchgott (Ed.), *Pharmacological and biophysical agents and behavior*. New York: Academic Press, 1971.
- Chase, T. N. Clinical studies of dopaminergic mechanisms. In E. Usdin (Ed.), *Neuropsychopharmacology of monoamines and their regulatory enzymes*. New York: Raven Press, 1974.
- Corrodi, H., Fuxe, K., & Lidbrink, D. Interaction between cholinergic and catecholaminergic neurones in rat brain. *Brain Research*, 1972, 43, 397-416.
- Costa, E., & Groppetti, A. Biosynthesis and storage of catecholamines in tissues of rats injected with various doses of d-amphetamine. In E. Costa & S. Garattini (Eds.), *Amphetamines and related compounds*. New York: Raven Press, 1970.
- Costall, B., & Naylor, R. J. Possible involvement of noradrenergic area of the amygdala with stereotyped behavior. *Life Sciences*, 1972, 11, 1135-1146.
- Costall, B., & Naylor, R. J. The importance of the ascending dopaminergic systems to the extrapyramidal and mesolimbic brain areas for the cataleptic action of the neuroleptic and cholinergic agents. *Neuropharmacology*, 1974, 13, 353-364.
- Costall, B., & Olley, J. E. Cholinergic and neuroleptic induced catalepsy: Modification by lesions in the globus pallidus and substantia nigra. *Neuropharmacology*, 1971, 10, 581-594.
- Desiderato, O., MacKinnon, J. R., & Hissom, H. Development of gastric ulcers in rats following stress termination. *Journal of Comparative and Physiological Psychology*, 1974, 87, 208-214.
- de Toledo, L., & Black, A. H. Retention of aversively motivated responses in rats. *Journal of Comparative and Physiological Psychology*, 1970, 71, 276-282.
- Ebel, A., Hermetet, J. C., & Mandel, P. Comparative study of acetylcholinesterase and choline acetyltransferase enzyme activity in brain of DBA and C57 mice. *Nature, New Biology*, 1973, 242, 56-57.
- Fibiger, H. C., Lynch, G. S., & Cooper, H. P. A biphasic action of central cholinergic stimulation on behavioral arousal in the rat. *Psychopharmacologia*, 1971, 20, 366-382.
- Fibiger, H. C., Phillips, A. G., & Zis, P. S. Deficits in instrumental responding after 6-hydroxy-dopamine lesions of the nigro-neostriatal dopaminergic projection. *Pharmacology Biochemistry and Behavior*, 1974, 2, 87-96.
- Fjalland, B., Moller-Nielsen, I. Methylphenidate antagonism of Haloperidol, interaction with cholinergic and anticholinergic drugs. *Psychopharmacologia*, 1974, 34, 111-118.
- Fuxe, K., & Ungerstedt, U. Histochemical, biochemical and functional studies on central monoamine neurons after acute and chronic amphetamine administration. In E. Costa & S. Garattini (Eds.), *Amphetamines and related compounds*. New York: Raven Press, 1970.
- Geller, A., Jarvik, M. E., & Robustelli, F. Incubation and the Kamin effect. *Journal of Experimental Psychology*, 1970, 85, 61-65.
- Gellhorn, E. *Autonomic imbalance and the hypothalamus*. Minneapolis: University of Minnesota Press, 1957.
- Grossman, S. P. Effect of chemical stimulation of the septal area on motivation. *Journal of Comparative and Physiological Psychology*, 1964, 58, 194-200.
- Herman, Z. S., Kmiecik-Kolada, K., Slominska-Zurek, J., & Szkilnik, R. Central effects of acetylcholine. *Psychopharmacologia*, 1972, 27, 223-232.
- Hitzeman, R. J., Loh, H. H., & Domino, E. F. Effect of scopolamine on the cerebral accumulation of ^{14}C -catecholamines from ^{14}C -tyrosine. *Pharmacology*, 1972, 8, 291-299.
- Ho, A. K. S., Singer, G., & Gershon, S. Biochemical evidence of adrenergic interaction with cholinergic function in the central nervous system of the rat. *Psychopharmacologia*, 1971, 21, 238-246.
- Irwin, S., Banuazizi, A., Kalsner, S., & Curtis, A. One trial learning in the mouse. I. Its characteristics and modification by experimental-seasonal variables. *Psychopharmacologia*, 1968, 12, 286-302.
- Izquierdo, I. Possible peripheral adrenergic and cholinergic mechanisms in pseudo-conditioning. *Psychopharmacologia*, 1974, 35, 189-193.
- Janowsky, D. S., El-Yousef, M. K., Davis, J. M., & Sekerke, H. J. Cholinergic antagonism of methylphenidate-induced stereotyped behavior. *Psychopharmacologia*, 1972, 27, 295-303.
- Javoy, F., Agid, Y., Bouvet, D., & Glowinski, J. Changes in neostriatal DA metabolism after carbachol or atropine microinjections into the substantia nigra. *Brain Research*, 1974, 68, 253-260.
- Javoy, F., Thierry, A. M., Kety, S. S., & Glowinski, J. The effect of amphetamine on the turn-

- over of brain norepinephrine in normal and stressed rats. *Communications in Behavioral Biology*, 1968, 1A, 43-48.
- Kamin, L. J. Retention of an incompletely learned avoidance response: Some further analyses. *Journal of Comparative and Physiological Psychology*, 1963, 56, 713-718.
- Karczmar, A. G., Scudder, C. L., & Richardson, D. L. Interdisciplinary approach to the study of behavior in related mice types. In S. Ehrenpreis & I. J. Kopin (Eds.), *Chemical approaches to brain function*. New York: Academic Press, 1973.
- Kasper-Pandi, P., Handing, R., & Usher, D. The effects of dexamethasone blockade of ACTH release on avoidance learning. *Physiology & Behavior*, 1970, 5, 361-365.
- Kato, L., Gozsy, B., Roy, P. B., & Groh, V. Histamine, serotonin, epinephrine and norepinephrine in the rat brain following convulsions. *International Journal of Neuropsychiatry*, 1967, 3, 46.
- Katzev, R. D., & Mills, S. K. Strain differences in avoidance conditioning as a function of the classical CS-US contingency. *Journal of Comparative and Physiological Psychology*, 1974, 87, 661-671.
- Kazic, T. Norepinephrine synthesis and turnover in the brain: Acceleration by physostigmine. In E. Usdin & S. H. Snyder (Eds.), *Frontiers in catecholamine research*. Oxford, England: Pergamon Press, 1973.
- Kelsey, J. E., & Grossman, S. P. Cholinergic blockade and lesion of the ventro-medial septum of the rat. *Physiology & Behavior*, 1969, 4, 837-845.
- Klein, S. B. Adrenal-pituitary influence in reactivation of avoidance learning memory in the rat after intermediate intervals. *Journal of Comparative and Physiological Psychology*, 1972, 79, 341-359.
- Klein, S. B., & Spear, N. E. Forgetting by the rat after intermediate intervals (Kamin effect) as retrieval failure. *Journal of Comparative and Physiological Psychology*, 1970, 71, 165-170. (a)
- Klein, S. B., & Spear, N. E. Reactivation of avoidance learning memory in the rat after intermediate retention intervals. *Journal of Comparative and Physiological Psychology*, 1970, 72, 498-504. (b)
- Klein, S. B., & Spear, N. E. Sequential variables and the Kamin effect. *Learning and Motivation*, 1973, 4, 357-365.
- Korf, J., Aghajanian, G. K., & Roth, R. H. Increased turnover of norepinephrine in the rat cerebral cortex during stress: Role of locus coeruleus. *Neuropharmacology*, 1973, 12, 933-938.
- Krieckhaus, E. E., Miller, N. E., & Zimmerman, P. Reduction of freezing behavior and improvement of shock avoidance by d-amphetamine. *Journal of Comparative and Physiological Psychology*, 1965, 60, 36-40.
- Kumar, R. Incubation of fear: Experiments on the "Kamin effect" in rats. *Journal of Comparative and Physiological Psychology*, 1970, 70, 258-263.
- Ladinsky, H., Consolo, S., Bianchi, S., Samanin, R., & Ghezzi, D. Cholinergic-dopaminergic interaction in the striatum: The effect of 6-hydroxydopamine or pimozide treatment on the increased striatal acetylcholine levels induced by apomorphine, piribedil and d-amphetamine. *Brain Research*, 1975, 84, 221-226.
- Levine, S., & Brush, F. R. Adrenocortical activity and avoidance learning as a function of time after avoidance training. *Physiology & Behavior*, 1967, 2, 385-388.
- Lewis, P. R., Shute, C. C. D., & Silver, A. Confirmation from choline acetylase analyses of a massive cholinergic innervation to the rat hippocampus. *Journal of Physiology*, 1967, 191, 215-224.
- Lloyd, K. G., & Hornykiewicz, O. Dopamine and other monoamines in the basal ganglia: Relation to brain dysfunctions. In P. Seeman & G. M. Brown (Eds.), *Frontiers in neurology and neuroscience research*. Toronto, Canada: University of Toronto Press, 1974.
- Lloyd, K. G., Stadler, H., & Bartholini, G. Dopamine and acetylcholine neurones in striatal and limbic structures: Effect of neuroleptic drugs. In E. Usdin & S. H. Snyder (Eds.), *Frontiers in catecholamine research*. Oxford, England: Pergamon Press, 1973.
- Longo, V. G. Behavioral and electroencephalographic effect of atropine and related compounds. *Pharmacological Review*, 1966, 18, 965-996.
- Lorens, S. Raphe lesions in cats: Forebrain serotonin avoidance behavior. *Pharmacology Biochemistry and Behavior*, 1973, 1, 487-490.
- Lynch, G., Mathews, D. A., Mosko, S., Parks, T., & Cotman, C. Induced acetylcholinesterase-rich layer in rat dentate gyrus following entorhinal lesions. *Brain Research*, 1972, 42, 311-318.
- Maier, S. F., Seligman, M. E. P., & Solomon, R. L. Pavlovian fear conditioning and learned helplessness: Effects on escape and avoidance behavior of (a) the CS-US contingency and (b) the independence of the US and voluntary responding. In B. A. Campbell & R. M. Church (Eds.), *Punishment and aversive behavior*. New York: Appleton-Century-Crofts, 1969.
- Mandell, A. J., & Knapp, S. The effects of chronic administration of some cholinergic and adrenergic drugs on the activity of choline acetyltransferase in the optic lobes of the chick brain. *Neuropharmacology*, 1971, 10, 513-516.
- Manto, P. G. Blockade of epinephrine-induced decrement in activity by scopolamine. *Psychonomic Science*, 1967, 7, 203-204.
- Manto, P. G. *An investigation of feedback mechanisms in stress-induced changes of autonomic balance*. Unpublished doctoral dissertation, Rutgers University, 1968.
- Maynert, E. W., & Levi, R. Stress-induced release of brain norepinephrine and its inhibition

- by drugs. *Journal of Pharmacology and Experimental Therapeutics*, 1964, 143, 90-95.
- McGeer, P. L., Grewaal, D. S., & McGeer, E. G. Influence of noncholinergic drugs on rat striatal acetylcholine levels. *Brain Research*, 1974, 80, 211-217.
- McMichael, J. S. Incubation of anxiety and instrumental behavior. *Journal of Comparative and Physiological Psychology*, 1966, 61, 208-211.
- Mellgren, S. I., & Srebro, B. Changes in acetylcholinesterase and distribution of degenerating fibers in the hippocampal region after septal lesions in the rat. *Brain Research*, 1973, 52, 19-36.
- Miller, W. R., & Seligman, M. E. P. Depression and the perception of reinforcement. *Journal of Abnormal Psychology*, 1973, 82, 62-73.
- Moore, K. E., & Lariviere, E. W. Effects of d-amphetamine and restraint on the content of biorepinephrine and dopamine in rat brain. *Biochemical Pharmacology*, 1963, 12, 1283-1288.
- O'Keefe, R., Sharman, D. F., & Vogt, M. Effect of drugs used in psychosis on cerebral dopamine metabolism. *British Journal of Pharmacology*, 1970, 38, 287-304.
- Oliverio, A. Contrasting effects of scopolamine on mice trained simultaneously with two different schedules of avoidance conditioning. *Psychopharmacologia*, 1967, 11, 39-51.
- Paulsen, E. C., & Hess, S. M. The rate of synthesis of catecholamines following depletion in guinea pig brain and heart. *Journal of Neurochemistry*, 1963, 10, 453.
- Pedersen, V. Potentiation of apomorphine effect (compulsive gnawing behavior) in mice. *Acta Pharmacologica et Toxicologica*, 1967, 25, 63.
- Pinel, J. P. J., Corcoran, M. E., & Malsbury, C. W. Incubation effect in rats: Decline of footshock produced activation. *Journal of Comparative and Physiological Psychology*, 1971, 77, 271-276.
- Pinel, J. P. J., & Mucha, R. F. Activity and reactivity in rats at various intervals after footshock. *Canadian Journal of Psychology*, 1973, 27, 112-118. (a)
- Pinel, J. P. J., & Mucha, R. F. Incubation and the Kamin effect in rats: Changes in activity and reactivity after footshock. *Journal of Comparative and Physiological Psychology*, 1973, 84, 661-667. (b)
- Pryor, G. T., Schlesinger, K., & Calhoun, W. H. Differences in brain enzymes among five inbred strains of mice. *Life Science*, 1966, 5, 2105-2111.
- Randrup, A., & Munkvad, I. Behavioral stereotypes induced by pharmacological agents. *Pharmakopsychiatric/Neuropsychopharmakologic*, 1968, 1, 18-26.
- Randrup, A., & Munkvad, I. Biochemical, anatomical and psychological investigations of stereotyped behavior induced by amphetamines. In E. Costa & S. Garattini (Eds.), *Amphetamines and related compounds*. New York: Raven Press, 1970.
- Rech, R. H., & Stolk, J. M. Amphetamine-drug interactions that relate brain catecholamines to behavior. In E. Costa & S. Garattini (Eds.), *Amphetamines and related compounds*. New York: Raven Press, 1970.
- Remington, G., & Anisman, H. Disruptive effects of epinephrine on active avoidance behavior: Alteration by scopolamine and d-amphetamine. *Pharmacology Biochemistry and Behavior*, 1974, 2, 427-430.
- Rescorla, R. A., & Solomon, R. L. Two process learning theory: Relationships between Pavlovian and instrumental learning. *Psychological Review*, 1967, 74, 151-182.
- Richardson, J. S. Basic concepts of psychopharmacological research as applied to the psychopharmacological analysis of the amygdala. *Acta Neurobiologiae Experimentalis*, 1974, 34, 543-562.
- Riley, E. P., Dunlap, W. P., & Hughes, L. F. The Kamin effect as a function of number of prior unsignalled inescapable shocks. *Journal of Comparative and Physiological Psychology*, 1975, 88, 565-569.
- Robustelli, F., Geller, A., & Jarvik, M. E. J. Biphasicity of the incubation curve. *Psychonomic Science*, 1970, 20, 129-130.
- Robustelli, F., Geller, A., & Jarvik, M. E. Biphasicity of the time-response curve and amnesic effect of electroconvulsive shock in passive avoidance conditioning. *Animal Learning and Behavior*, 1973, 1, 251-253.
- Rosecrans, J. A. Brain amine changes in stressed and normal rats pretreated with various drugs. *Archives of International Pharmacodynamics*, 1969, 180, 460-470.
- Rosic, N., & Bignami, G. Depression of two-way avoidance learning and enhancement of passive avoidance learning by small doses of physostigmine. *Neuropharmacology*, 1970, 9, 311-316.
- Rosic, N., Bignami, G., & Gatti, G. L. The use of shuttle-box avoidance in the study of anticholinesterase effects and of anticholinesterase-oxime interactions. In S. B. Baker (Ed.), *The correlation of adverse effects in man with observations in animals*. Amsterdam: Excerpta Medica, 1971.
- Scapagnini, U., Annunziato, L., & Preziosi, P. Role of brain norepinephrine in stress regulation. In S. Nemeth (Ed.), *Hormones, metabolism and stress: Recent progress and perspectives*. Bratislava, Czechoslovakia: Publishing House of the Slovak Academy of Sciences, 1973.
- Scheel-Kruger, J. Central effects of anticholinergic drugs measured by the apomorphine gnawing test in mice. *Acta Pharmacologica et Toxicologica*, 1970, 28, 1-16.
- Schildkraut, J. J. The catecholamine hypothesis of affective disorders—a review of supporting evidence. *American Journal of Psychiatry*, 1965, 122, 502-522.
- Seligman, M. E. P. On the generality of the laws of learning. *Psychological Review*, 1970, 77, 406-418.
- Seligman, M. E. P. Depression and learned helplessness. In R. J. Freedman & M. M. Katz (Eds.), *The psychology of depression: Contemporary*

- porary theory and research. Washington, D.C.: V. H. Winston, 1974.
- Seligman, M. E. P., Klein, D. C., & Miller, W. R. Depression. In H. Leitenberg (Ed.), *Handbook of behavior therapy*. New York: Appleton-Century-Crofts, in press.
- Seligman, M. E. P., Maier, S. F., & Solomon, R. L. Unpredictable and uncontrollable aversive events. In F. R. Brush (Ed.), *Aversive conditioning and learning*. New York: Academic Press, 1971.
- Sethy, V. H., & Van Woert, M. H. Effect of L-Dopa on brain acetylcholine and choline in rats. *Neuropharmacology*, 1973, 12, 27-31.
- Sethy, V. H., & Van Woert, M. H. Modification of striatal acetylcholine concentration by dopamine receptor agonists and antagonists. *Research Communications in Chemical Pathology and Pharmacology*, 1974, 8, 13-28.
- Shurman, A. J., & Katzev, R. D. Escape/avoidance responding in rats depends on strain and number of inescapable preshocks. *Journal of Comparative and Physiological Psychology*, 1975, 88, 548-553.
- Singh, P. J., Sakellaris, P. C., & Brush, F. R. Retention of active and passive avoidance responses tested in extinction. *Learning and Motivation*, 1971, 2, 305-323.
- Snider, N., Marquis, H. A., Black, M., & Suboski, M. D. Adrenal corticosteroids and the Kamin effect. *Psychonomic Science*, 1971, 22, 309-310.
- Spear, N. E. Retrieval of memory in animals. *Psychological Review*, 1973, 80, 163-194.
- Spear, N. E., Klein, S. B., & Riley, E. P. The Kamin effect as state dependent learning: Memory retrieval failure in the rat. *Journal of Comparative and Physiological Psychology*, 1971, 74, 416-425.
- Stadler, H., Lloyd, K. G., & Bartholini, G. Dopaminergic inhibition of striatal cholinergic neurons: Synergistic blocking action of γ -butyrolactone on neuroleptic drugs. *Naunyn-Schmiedeberg's Archives of Pharmacology*, 1974, 283, 129-134.
- Stadler, H., Lloyd, K. G., Gadea-Ciria, M., & Bartholini, G. Enhanced striatal acetylcholine release by chlorpromazine and its reversal by apomorphine. *Brain Research*, 1973, 55, 476-480.
- Stein, L., & Wise, C. D. Serotonin and behavioral inhibition. *Advances in Biochemical Pharmacology*, 1974, 11, 281-291.
- Steranka, L. R., & Barrett, R. J. Kamin effect in rats: Differential retention or differential acquisition of an active-avoidance response. *Journal of Comparative and Physiological Psychology*, 1973, 85, 324-330.
- Steranka, L. J., & Barrett, R. J. Facilitation of avoidance acquisition by lesion of the median raphe nucleus: Evidence for serotonin as a mediator of shock-induced suppression. *Behavioral Biology*, 1974, 11, 205-215.
- Stolk, J. M., Conner, R. L., Levine, S., & Barchas, J. D. Brain norepinephrine metabolism and shock-induced fighting behavior in rats: Differential effects of shock and fighting on the neurochemical response to a common footshock stimulus. *Journal of Pharmacology and Experimental Therapeutics*, 1974, 190, 193-209.
- Stone, E. A. Swim-stress-induced inactivity: Relation to body temperature and brain norepinephrine, and effect of d-amphetamine. *Psychosomatic Medicine*, 1970, 32, 51-59.
- Stromberg, U. Dopa effects on motility in mice: Potentiation by MK485 and dexchlorpheniramine. *Psychopharmacologia*, 1970, 18, 58-67.
- Suboski, M. D., Marquis, H. A., Black, M., & Platenius, P. Adrenal and amygdala function in the incubation of aversively conditioned responses. *Physiology & Behavior*, 1970, 5, 283-289.
- Suits, E., & Isaacson, R. L. The effects of scopolamine hydrobromide on one-way and two-way avoidance learning in rats. *International Journal of Neuropharmacology*, 1968, 7, 441-446.
- Swonger, A. K., & Rech, R. H. Serotonergic and cholinergic involvement in habituation of activity and spontaneous alternation of rats in a Y-maze. *Journal of Comparative and Physiological Psychology*, 1972, 81, 509-522.
- Tarpy, R. M. Incubation of anxiety as measured by response suppression. *Psychonomic Science*, 1966, 4, 189-190.
- Thierry, A. M. Effects of stress on the metabolism of serotonin and norepinephrine in the central nervous system of the rat. In S. Nemeth (Ed.), *Hormones, metabolism and stress: Recent progress and perspectives*. Bratislava, Czechoslovakia: Publishing House of the Slovak Academy of Sciences, 1973.
- Thierry, A., Javoy, F., Glowinski, J., & Kety, S. S. Effects of stress on the metabolism of norepinephrine, dopamine and serotonin in the central nervous system of the rat. 1. Modification of norepinephrine turnover. *Journal of Pharmacology and Experimental Therapeutics*, 1968, 163, 163-171.
- Thomas, E., Balter, A. Learned helplessness: Amelioration of symptoms by cholinergic blockade of the septum. *Science*, in press.
- Thompson, R. W., & Neilson, C. The effect of scopolamine on the Kamin effect: A test of the parasympathetic overreaction hypothesis. *Psychonomic Science*, 1972, 28, 140-142.
- Thornburg, J. E., & Moore, K. E. Inhibition of anticholinergic drug-induced locomotor stimulation in mice by α -methyltyrosine. *Neuropharmacology*, 1973, 12, 1179-1186.
- Weiss, J. M. Effects of coping behavior in different warning signal conditions on stress pathology in rats. *Journal of Comparative and Physiological Psychology*, 1971, 77, 1-13. (a)
- Weiss, J. M. Effects of coping behavior with and without a feedback signal on stress pathology in rats. *Journal of Comparative and Physiological Psychology*, 1971, 77, 22-30. (b)
- Weiss, J. M. Effects of punishing the coping response (conflict) on stress pathology in rats.

- Journal of Comparative and Physiological Psychology*, 1971, 77, 14-21. (c)
- Weiss, J. M., Glazer, H. I., & Pohorecky, L. A. Coping behavior and neurochemical changes: An alternative explanation for the original "learned helplessness" experiments. In *Relevance of the psychopathological animal model to the human*. New York: Plenum Press, in press.
- Weiss, J. M., Kreickhaus, E. E., & Conte, R. Effects of fear conditioning on subsequent avoidance behavior and movement. *Journal of Comparative and Physiological Psychology*, 1968, 65, 413-421.
- Weiss, J. M., Stone, E. A., & Harrell, N. Coping behavior and brain norepinephrine level in rats. *Journal of Comparative and Physiological Psychology*, 1970, 72, 153-160.
- Welch, A. S., & Welch, B. L. Effect of stress and para-chlorophenylalanine upon brain serotonin, 5-hydroxyindoleacetic acid and catecholamines in grouped and isolated mice. *Biochemical Pharmacology*, 1968, 17, 699.
- Welch, B. L., & Welch, A. S. Control of brain catecholamines and serotonin during acute stress and after d-amphetamine by natural inhibition of monoamine oxidase: An hypothesis. In E. Costa & S. Garattini (Eds.), *Amphetamines and related compounds*. New York: Raven Press, 1970.
- Winokur, G. The types of affective disorders. *Journal of Nervous and Mental Disease*, 1973, 156, 82-96.
- Zammit-Montebello, A., Black, M., Marquis, H. A., & Suboski, M. D. Incubation of passive avoidance in rats: Shock intensity and pretraining. *Journal of Comparative and Physiological Psychology*, 1969, 69, 579-582.

(Received July 23, 1974; revision received June 18, 1975)

LETTERS TO THE EDITOR

Neural Basis of Shape Constancy

Milner's article on visual shape recognition is valuable for spelling out plausible possibilities of feature analyzing circuits that are quite amenable to experimental test (Milner, 1974). However, his model does not cover the important question, less accessible to present physiological techniques, of the spatial relations of the features internal to a pattern.

Although certain details must be discarded to achieve perceptual constancy, these feature analyzers lose critical information. For example, the distance ratio cell preserves information about the ratio but not about which line is which. It is not clear whether the mechanism that restores information about locus, for purposes of directing responses, also works for internal structure. If the net effect of the nonsynchronized firing of cells that perceive different parts of a display is simply to convert a question of spatial organization into one of temporal organization, then the basic problem still is not solved.

It seems doubtful that feature analysis in sensory systems continues on at higher levels toward single, "pontifical" neurons, prepared for every possible experience (Perkel & Bullock, 1968, pp. 284-286). Alternatively, higher levels of analysis might involve mapping of spatial stimulus information into some sort of organized pattern of multineuronal interaction (e.g., Pribram, 1971), aided by feature analysis properties of single neurons.

Brain activity is visible through amplification channels only at a few points at a time. Until some fundamental advance in instrumentation occurs, we need inferential tools that allow us to recognize possible organized activity in neural networks.

REFERENCES

- Milner, P. M. A model for visual shape recognition. *Psychological Review*, 1974, 81, 521-535.
Perkel, D. H., & Bullock, T. H. Neural coding. *Neurosciences Research Program Bulletin*, 1968, 6, 221-348.

Pribram, K. H. *Languages of the brain*. Englewood Cliffs, N.J.: Prentice-Hall, 1971.

Robert B. Glassman
Department of Psychology
Lake Forest College
Lake Forest, Illinois 60045

(Received March 18, 1975)

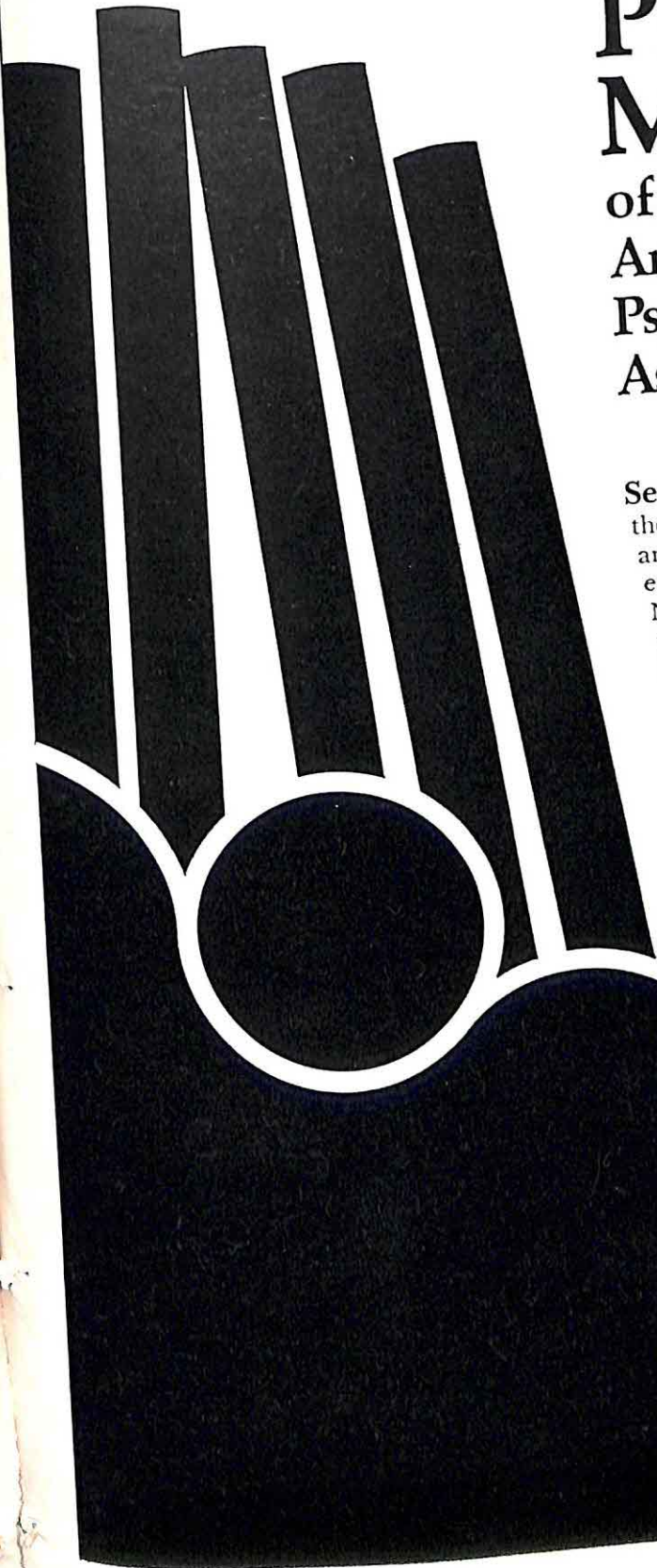
Reply by Milner

Glassman is correct in pointing out that the attention mechanism I described functions only for whole objects and cannot explain how, for example, one might focus attention on the top left-hand corner of a figure, or the brushwork on the grass in the foreground of a painting. I think the answer to this must lie in the complexity of the available concepts, especially abstract concepts such as "top," "second," "middle," and so on. If several such concepts could be activated simultaneously, it would result in levels of neural activity higher in certain parts of the visual field than in others and a consequent preferential use of neurons there for controlling responses.

It is possible that I did not adequately stress that the features resulting from the processing described in my paper converge again at a concept (or "cell-assembly") level. Glassman apparently considers it necessary to emphasize that relationship with his point regarding the unlikelihood of "pontifical" neurons. I entirely agree with theories of multineuronal representation of concepts, but I am not too optimistic that additional "inferential tools" can be substituted for improved technology.

Peter M. Milner
Department of Psychology
McGill University
Montreal, Quebec H3C 3G1

(Received April 7, 1975)



Publication Manual

of the
American
Psychological
Association

Second edition, 1974. Current guide to the preparation of manuscripts for APA and related journals. This edition adds many examples, including an entire sample paper. New sections on student papers, metrication, and mathematical copy. Enlarged sections on content and organization of an article, typing the manuscript, and all aspects of APA style, including reference lists and tables. 136 pp. \$4.

Please note that payment must accompany all orders amounting to \$15 or less except for those orders submitted on institutional purchase order forms.

American Psychological Association
Publication Sales Dept.
1200 17th Street, N.W.
Washington, D.C. 20036



Thesaurus of Psychological Index Terms

Contains over 4,000 terms representing an authoritative, unambiguous, and hierarchically structured vocabulary used by psychology and related disciplines. This vocabulary will facilitate precise content representation of concepts described in psychological literature and allow for more flexibility in automated search and retrieval operations. These two factors promote more effective communication of psychological concepts in original works and research, ultimately aiding the content of publications to be focused in accordance with the interest patterns of psychologists.

This **Thesaurus**, developed by the Psychological Abstracts Information Services, is divided into three sections:

The **relationship** section shows the relationships among concepts that are associated with each other. Term relationships noted are: synonymous, broader, narrower, or related.

The **alphabetical** section presents each preferred term in alphabetical order, enabling rapid identification of terms used for indexing.

The **hierarchical** section contains 17 broad content areas in which terms are presented in descending order according to breadth of concept represented. The 17 categories correspond to major classifications used in *Psychological Abstracts*.

The relationship, alphabetical, and hierarchical sections will assist the user in selecting appropriate terms for different uses.

Administrators

- determine areas of research for an organization
- assure representative expertise in personnel selection

Researchers and Authors

- find synonyms for a term representing a concept
- select terms that represent a concept
- define subject matter topics for a paper or book

Educators

- structure a curriculum
- use as a tutorial aid

Consultants

- search the literature
- aid in retrieval of information

Librarians

- provide access to psychological literature
- aid in cataloging psychology material
- aid in ordering new psychological literature

Students

- locate concepts relevant to psychology
- focus areas of study

The **Thesaurus** contains 350 pages with a hard-cover and costs \$12. Please note that all orders amounting to \$15 or less must be accompanied by full payment except for those orders submitted on institutional purchase order forms.

American Psychological Association
Publication Sales Department
1200 17th Street, N.W.
Washington, D.C. 20036



1972-74 Cumulative Indexes to Psychological Abstracts

These three-year **Indexes** contain all records published in the *Psychological Abstracts* from January 1972 through December 1974. Two kinds of indexes are available:

The **Subject Index** consolidates references to the world's literature in psychology and related disciplines that were summarized in 36 monthly issues of *Psychological Abstracts*. Index terms representing the major concepts described in this literature have been applied. Phrases describing subject matter content appear under each term, enabling the user to select those references which are relevant to their needs and interests. Entries in the index contain subject index terms, *see* and *see also* reference listings, descriptive phrases, and *PA* volume and record number (two volumes).

The **Author Index** contains references to authors who produced material that was summarized in 36 monthly issues of *Psychological Abstracts*. As many as four authors are referenced for each entry. Thus, all the publications of individual investigators who conduct programmatic research in collaboration with others can be easily identified. Primary entries in the index contain up to four author's names, title of source document, year of publication, volume and issue numbers, inclusive pagination, and *PA* volume and record numbers. Secondary authors are listed by name only, with a cross-reference to the primary author and *PA* volume and record number (one volume).

The **Cumulative Indexes** for 1972-74 utilize index terms from the new APA *Thesaurus of Psychological Index Terms*. The size of the **Indexes** remains compat-

ible with those published previously. Pages are 10 x 14 inches, printed on durable acid-free paper, and encased in Class A library binding.

	U.S. and other PAIS countries	Non PAIS countries
1972-74 Cumulative Subject Index to PA (two volumes)	\$380	\$390
1972-74 Cumulative Author Index to PA (one volume)	\$235	\$245
Both Indexes for 1972-74	\$485	\$495
1969-71 Cumulative Indexes to Psychological Abstracts are also available from APA.		
1969-71 Cumulative Subject Index to PA (two volumes)	\$360	\$390
1969-71 Cumulative Author Index to PA (one volume)	\$220	\$240
Both Indexes for 1969-71	\$465	\$495

Earlier editions of the **Cumulative Indexes**, for the period 1927-1968, may be obtained by writing to: G. K. Hall and Company, 70 Lincoln Street, Boston, Massachusetts 02111. Please do not write to APA to obtain information about these editions.

Ordering information. Please send your request for the 1972-74 and 1969-71 **Indexes** accompanied by full payment or institutional purchase order to:

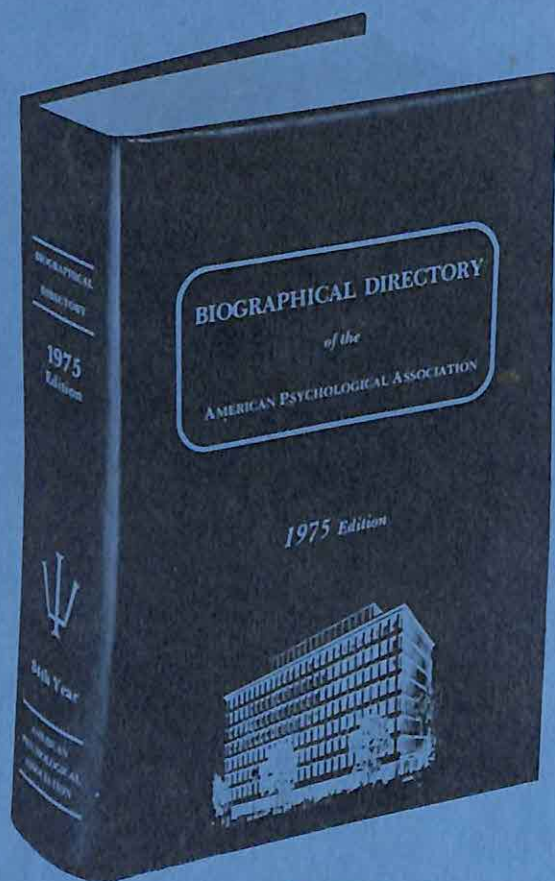
American Psychological Association
Publication Sales Department
1200 17th Street, N.W.
Washington, D.C. 20036



1975

Biographical Directory

of the American Psychological Association



- 1500 pages (estimated)
- Listing approximately 39,500 APA members
- \$20 to APA members; \$40 to nonmembers and organizations.

Ordering Information: Please send your request accompanied by full payment or institutional purchase order to:

American Psychological Association
Publication Sales Dept.
1200 17th Street, N.W.
Washington, D.C. 20036



This revised and expanded **Directory** lists name and address, telephone number, education, present major field, areas of specialization, licensure or certification as a psychologist, certification as a school psychologist, ABPP and ABPH diplomate status, and APA membership status. The **Directory** also contains a geographical index section and a membership roster for each APA division. APA Bylaws, present and past APA officers, Ethical Standards, and current data on laws governing the practice of psychology in the United States and Canada are included.

PSYCHOLOGICAL REVIEW

GEORGE MANDLER, EDITOR
JEAN M. MANDLER, ASSOCIATE EDITOR
UNIVERSITY OF CALIFORNIA, SAN DIEGO

Vol. 82, No. 6

November 1975

Articles

- The Endogenous-Exogenous Partition in Attribution Theory
Arie W. Kruglanski 387
- A Spreading-Activation Theory of Semantic Processing
Allan M. Collins and Elizabeth F. Loftus 407
- Subjective Duration and Psychophysics
Hannes Eisler 429
- The Two-Transient (Masking) Paradigm
Ethel Matin 451
- On the Role of Context Effects in Psychophysical Judgment
Norman H. Anderson 462

Theoretical Note

- Parallel Psychometric Functions from a Set of Independent Detectors
David M. Green and R. Duncan Luce 483

406

Announcement

This is the last issue of Volume 82.
Title page and contents of the volume appear herein.

Published bimonthly by the American Psychological Association, Inc.

Editorial Consultants

The editor gratefully acknowledges the efforts of the following consultants who helped review manuscripts for this issue. They are part of a large group of psychologists who have kindly agreed to assist in the editing of manuscripts for *Psychological Review*.

ROBERT M. BOYNTON
PETER L. CARLTON
HERBERT H. CLARK
CLYDE H. COOMBS
LYNN A. COOPER
MICHAEL E. DAWSON
TRYGG ENGLEN
ROBERT L. FANTZ
WALTER C. GOGEL
WILLIAM C. GORDON
B. M. HERMELIN

JAMES KAHAN
HAROLD H. KELLEY
A. B. KRISTOFFERSON
DAVID LABERGE
R. DUNCAN LUCE
DONALD I. A. MACLEOD
DAVID MCNEILL
D. E. MEYER
D. E. MITCHELL
ALLEN PARDUCCI
W. F. PROKASY

ELEANOR ROSCH
PETER H. SCHILLER
MARILYN SHATZ
NORMAN SPEAR
JOHN E. R. STADDON
ROBERT TEGHTSOONIAN
JOHN THEIOS
MICHEL TREISMAN
PAUL VITZ
ROGER A. WEBB
NAOMI WEISSTEIN

The *Psychological Review* publishes articles that make theoretical contributions to any area of scientific psychology. Preference is given to papers that advance theory rather than review it, and to statements that are specifically theoretical rather than programmatic. Surveys of the literature, problems of method and design, and papers that primarily report empirical findings are ordinarily not appropriate. Letters not exceeding 250 words on topics of theoretical interest should be submitted in duplicate.

Preparation of Manuscripts: Instructions for the preparation of manuscripts appear in the *Publication Manual of the American Psychological Association* (2nd ed.). Submit three copies; one copy must be the original typed copy. All parts of the manuscript must be double spaced. In order to facilitate blind reviewing, the author's name should not appear on the manuscript or abstract, but each copy of the manuscript should include a separate title page showing the title and the author's name and affiliation. Footnotes that identify the author should be typed on separate pages. Figures should be submitted as glossy prints ready for reproduction; duplicate figures may be photocopies. Metric equivalents must be given for all nonmetric measurements. Authors should keep a copy of the manuscript to guard against loss.

Abstracts: Manuscripts must be accompanied by an abstract of 100 to 175 words typed on a separate sheet of paper. Detailed instructions for preparation of abstracts appear in the *Publication Manual* or may be obtained from the Editor.

Manuscripts: Submit manuscripts in triplicate to the Editor, GEORGE MANDLER, Department of Psychology, University of California, San Diego, La Jolla, California 92037 according to instructions provided above.

Copyright and Permissions: All rights reserved. Written permission must be obtained from the American Psychological Association for copying or reprinting text of more than 500 words, tables, or figures. Permissions are normally granted contingent upon like permission of the author, inclusion of the APA copyright notice on the first page of reproduced material, and payment of a fee of \$10 per page of APA material. Permission fees are waived for authors who wish to reproduce their own articles. Address requests for permission to the Permissions Office, American Psychological Association, 1200 Seventeenth Street, N.W., Washington, D. C. 20036.

Subscriptions: Subscriptions are available on a calendar year basis only (January through December). Nonmember rates for 1975: \$15 domestic, \$16 foreign, \$3 single issue. APA member rate: \$6. Write to Subscription Section, APA.

Reprints: Authors may order reprints of their articles from the printer when they receive galley proofs.

Back Issues and Back Volumes: For information regarding back issues or back volumes write to Order Department, APA.

Microform Editions: For information regarding microform editions write to any of the following: Johnson Associates, Inc., P.O. Box 1017, Greenwich, Connecticut 06830; University Microfilms, Ann Arbor, Michigan 48106; or Princeton Microfilms, Princeton, New Jersey 08540.

Change of Address: Send change of address notice and a recent mailing label to the attention of the Subscription Section, APA 30 days prior to the actual change of address. Undelivered copies resulting from address changes will not be replaced; second-class mail is not forwarded unless subscribers notify the local post office in writing that they will guarantee second-class forwarding postage.

Published bimonthly (beginning in January) in one volume per year, six issues, by the American Psychological Association, Inc., 1200 Seventeenth Street, N.W., Washington, D. C. 20036. Printed in the U.S.A. Second-class postage paid at Washington, D. C.

Journal Staff

ANITA DEVIVO
Executive Editor

ARTHUR HAUPT
Supervising Technical Editor

KENDALL C. PATTERSON
Technical Editor

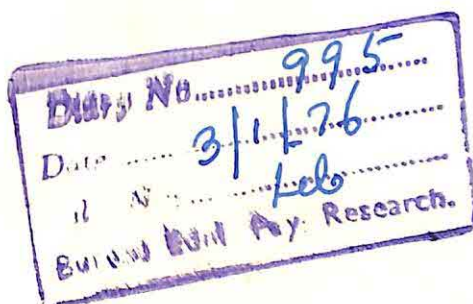
EDWARD B. HARNER
Advertising Manager

ANNE REDMAN
Subscription Manager

PSYCHOLOGICAL REVIEW

GEORGE MANDLER, EDITOR
JEAN M. MANDLER, ASSOCIATE EDITOR
UNIVERSITY OF CALIFORNIA, SAN DIEGO

VOLUME 82
1975



Published bimonthly by the
American Psychological Association, Inc.
1200 Seventeenth Street, N.W., Washington, D.C. 20036
Copyright © 1975 by the American Psychological Association, Inc.

Editorial Consultants

NORMAN ADLER
WESLEY ALLINSMITH
IRWIN ALTMAN
JOHN R. ANDERSON
NORMAN H. ANDERSON
J. W. ATKINSON

JACOB BECK
LILLIAN BELMONT
DARYL BEM
L. BERKOWITZ
INA M. BILODEAU
ROBERT A. BJORK
LOIS BLOOM
ROBERT BOICE
ROBERT C. BOLLES
GORDON H. BOWER
KENNETH S. BOWERS
ROBERT M. BOYNTON
BRUNO BREITMEYER
JOHN L. BROWN
M. PHILIP BRYDEN
ALEXANDER M. BUCHWALD

JOSEPH CAMPIONE
PETER L. CARLTON
DORWIN CARTWRIGHT
ROSALIND D. CARTWRIGHT
COURTNEY CAZDEN
LOREN CHAPMAN
J. M. CHERTKOFF
HERBERT H. CLARK
GEORGE COLLIER
A. M. COLLINS
BARRY E. COLLINS
CLYDE H. COOMBS
LYNN A. COOPER
VIRGINIA C. CRANDALL
CHARLES CRAWFORD

JAMES H. DAVIS
GERALD DAVISON
ROBYN DAWES
MICHAEL E. DAWSON

TRYGG ENGEN

ROBERT L. FANTZ
JOAN E. FOLEY

C. R. GALLISTEL
WALTER C. GOGEL
PAUL E. GOLD
WILLIAM C. GORDON
NORMAN GRAHAM

MARSHALL HAITH
O. J. HARVEY
JOHN R. HAYES
B. M. HERMELIN
STEPHEN A. HILLYARD

DOUGLAS HINTZMAN
JERRY HIRSCH
JULIAN HOCHBERG
LEO M. HURVICH

CHESTER A. INSKO
FRANCIS W. IRWIN

RICHARD JAGACINSKI
IRVING L. JANIS
H. M. JENKINS
EDWARD E. JONES

JAMES KAHAN
DANIEL KATZ
HAROLD H. KELLEY
DEBORAH KEMMLER
MARCEL KINSBOURNE
SAMUEL S. KOMORITA
DAVID H. KRANTZ
A. B. KRISTOFFERSON

DAVID LABERGE
BIBB LATANÉ
MARK R. LEPPER
HOWARD LEVANTHAL
GERALD S. LEVENTHAL
GEORGE LEVINGER
SARAH LICHTENSTEIN
LEAH L. LIGHT
R. DUNCAN LUCE
DAVID T. LYKKEN

DONALD I. A. MACLEOD
SALVATORE R. MADDI
STEPHEN A. MADIGAN
JAMES MCGAUGH
DAVID MCNEILL
D. E. MEYER
JEFFREY MILLER
RALPH MILLER
D. E. MITCHELL
JOHN MORTON
BENNETT MURDOCK, JR.

LLOYD NAKATANI
ALLEN NEWELL
ELISSA NEWPORT
KENT L. NORMAN

PAUL A. OBRIST

ROBERT G. PACHELLA
ALLEN PARDUCCI
ANNE D. PICK
ROY W. PICKENS
DAVID PREMACK
W. F. PROKASY

D. A. RILEY
ELLIOT ROGERS

ELEANOR ROSCH
SEYMOUR ROSENBERG
ERNST ROTHKOPF
DAVID E. RUMELHART

THEODORE SARBIN
EVELYN SATINOFF
PETER H. SCHILLER
KURT SCHLESINGER
ALLEN M. SCHNEIDER
DAVID J. SCHNEIDER
CARMI SCHOOLER
MARILYN SHATZ
RICHARD M. SHIFFRIN
JOAN G. SNODGRASS
NORMAN SPEAR
HARRY G. SPERLING
LARRY R. SQUIRE
JOHN E. R. STADDON
SAUL STERNBERG
DANIEL STOKOLS
CHARLES F. STOMEYER III
HANS H. STRUPP

JOHN E. TAPLIN
ROBERT TEGHTSOONIAN
PHILIP E. TEITELBAUM
DAVIDA TELLER
JOHN THEIOS
EARL THOMAS
HOBEN THOMAS
RICHARD F. THOMPSON
JAMES T. TOWNSEND
MICHEL TREISMAN

W. R. UTTAL

ELLIOT S. VALENSTEIN
STUART VALINS
PAUL VITZ

HANS WALLACH
ROC E. WALLEY
ROGER A. WEBB
WILSE B. WEBB
NAOMI WEISSTEIN
GERALD WESTHEIMER
RICHARD E. WHALEN
WAYNE A. WICKELGREN
THOMAS D. WICKENS
MORTON WIENER
FREDERICK WIGHTMAN
JERI D. WINE
JULIUS WISNER
GEORGE WOLFORD

AL YONAS

ROBERT B. ZAJONC
DAVID ZEAMAN
MICHAEL D. ZEILER

Journal Staff

Anita DeVivo
Executive Editor

Edward B. Harner
Advertising Manager

Kendall C. Patterson
Technical Editor

Arthur Haupt
Supervising Technical Editor

Anne Redman
Subscription Manager

Contents of Volume 82

Anderson, Norman H. On the role of context effects in psychophysical judgment	462
Anisman, Hymie. Time-dependent variations in aversively motivated behaviors: Non-associative effects of cholinergic and catecholaminergic activity	359
Carpenter, Patricia A., and Just, Marcel Adam. Sentence comprehension: A psycholinguistic processing model of verification	45
Collins, Allan M., and Loftus, Elizabeth F. A spreading-activation theory of semantic processing	407
Dienstbier, Richard A., Hillman, Donald, Lehnhoff, John, Hillman, Judith, and Valkenaar, Maureen C. An emotion-attribution approach to moral behavior: Interfacing cognitive and avoidance theories of moral development	299
Eisler, Hannes. Subjective duration and psychophysics	429
Folger, Robert. <i>See</i> Insko, Chester A.	
Glassman, Robert B. Neural basis of shape constancy	386
Green, David M., and Luce, R. Duncan. Parallel psychometric functions from a set of independent detectors	483
Grossman, Sebastian P. Role of the hypothalamus in the regulation of food and water intake	200
Hillman, Donald. <i>See</i> Dienstbier, Richard A.	
Hillman, Judith. <i>See</i> Dienstbier, Richard A.	
Hollan, James D. Features and semantic memory: Set-theoretic or network model?	154
Insko, Chester A., Worchel, Stephen, Folger, Robert, and Kutkus, Arūnas. A balance theory interpretation of dissonance	169
Just, Marcel Adam. <i>See</i> Carpenter, Patricia A.	
Kaestner, Elisabeth. <i>See</i> Woods, Stephen C.	89
Killeen, Peter. On the temporal control of behavior	1
Klinger, Eric. Consequences of commitment to and disengagement from incentives	387
Kruglanski, Arie W. The endogenous-exogenous partition in attribution theory	
Kutkus, Arūnas. <i>See</i> Insko, Chester A.	
Lehnhoff, John. <i>See</i> Dienstbier, Richard A.	344
LeVere, T. E. Neural stability, sparing, and behavioral recovery following brain damage	
Loftus, Elizabeth F. <i>See</i> Collins, Allan M.	
Luce, R. Duncan. <i>See</i> Green, David M.	
Mackintosh, N. J. A theory of attention: Variations in the associability of stimuli with reinforcement	276
Markus, Gregory B. <i>See</i> Zajonc, R. B.	150
Martin, Edwin. Generation-recognition theory and the encoding specificity principle	451
Martin, Ethel. The two-transient (masking) paradigm	386
McGuinness, Diane. <i>See</i> Pribram, Karl H.	
Milner, Peter M. Reply by Milner	
Ozog, Gregory. <i>See</i> Weisstein, Naomi.	
Panksepp, Jaak. Metabolic hormones and regulation of feeding: A reply to Woods, Decke, and Vasselli	158
Pribram, Karl H., and McGuinness, Diane. Arousal, activation, and effort in the control of attention	116
Rips, Lance J., Smith, Edward E., and Shoben, Edward J. Set-theoretic and network models reconsidered: A comment on Hollan's "Features and semantic memory"	156

Schmidt, Rainer. <i>See</i> Vorberg, Dirk.	
Schmidt, Richard A. A schema theory of discrete motor skill learning	225
Shoben, Edward J. <i>See</i> Rips, Lance J.	
Smith, Edward E. <i>See</i> Rips, Lance J.	
Stokols, Daniel. Toward a psychological theory of alienation	26
Szoc, Ronald. <i>See</i> Weisstein, Naomi.	
Tulving, Endel, and Watkins, Michael J. Structure of memory traces	261
Valkenaar, Maureen C. <i>See</i> Dienstbier, Richard A.	
Vasselli, Joseph R. <i>See</i> Woods, Stephen C.	
Vorberg, Dirk, and Schmidt, Rainer. A testable property of some discrete-state models for recognition memory	316
Watkins, Michael J. <i>See</i> Tulving, Endel.	
Weisstein, Naomi, Ozog, Gregory, and Szoc, Ronald. A comparison and elaboration of two models of metacontrast	325
Wolford, George. Perturbation model for letter identification	184
Woods, Stephen C., Kaestner, Elisabeth, and Vasselli, Joseph R. Insulin, growth hor- mone, body weight, and feeding: A reply to Panksepp	165
Worchel, Stephen. <i>See</i> Insko, Chester A.	
Zajonc, R. B., and Markus, Gregory B. Birth order and intellectual development	74

PSYCHOLOGICAL REVIEW

Copyright © 1975 by the American Psychological Association, Inc.

VOL. 82, No. 6

NOVEMBER 1975

The Endogenous-Exogenous Partition in Attribution Theory

Arie W. Kruglanski
Tel-Aviv University
Tel-Aviv, Israel

Within lay explanation of actions, several significant inferences are assumed to follow from the partition between endogenous and exogenous attributions. An action is said to be endogenously attributed when it is judged to constitute an end in itself and exogenously attributed when it is judged to serve as a means to some further end. A theoretical framework is developed whereby the action's endogenous attribution is linked with the inferences of intrinsic motivation, subjective freedom, and the action's underlying intention. The endogenous-exogenous distinction is proposed to replace the frequently invoked partition between the action's internal and external causes. Both conceptual and empirical considerations are put forth in favor of such a replacement. Classical attribution topics to which the internal-external partition has been applied are reinterpreted in terms of the endogenous-exogenous distinction, and novel data are reported that support the latter framework. Finally, several categories of conditions for endogenous (or exogenous) attributions are identified, and possible directions of further research within the endogenous-exogenous framework are suggested.

Attribution theory (Jones & Davis, 1965; Jones, Kanouse, Kelley, Nisbett, Valins, & Weiner, 1972; Kelley, 1967, 1971a, 1971b, 1973) deals with the explanation of effects via the identification of their causes. Frequently the causes of behaviors have been partitioned into those *internal* to the person and those *external* to him (internal causes have been alternatively labeled as personal and external as environmental). The internal-external division has been regarded

as basic to attribution processes (Kelley, 1967, p. 194) in the sense that the designation of a cause as internal has been assumed to identify it meaningfully so that explanation of the effect would readily follow. In particular the internal-external framework was assumed to yield inferences as diverse as those of (a) intrinsic motivation (see, e.g., deCharms, 1968; Deci, 1971), (b) subjective freedom (Kelley, 1967; Kruglanski & Cohen, 1973), (c) perceived pain (Nisbett & Valins, 1971), (d) positive attitudes (Bem, 1972), and (e) sincere reporting (Bem, 1972; Jones & Davis, 1965).

The present paper introduces the distinction between *endogenous* and *exogenous* attribution (analogous to the distinction between means and ends) to replace the internal-external partition in the lay explanation of actions. It is argued that the

I am indebted to Bonnie Erickson, Chet Insko, Ned Jones, Hal Kelley, Moshe Kroy, Andrea Sedlak, and John Thibaut for comments on an earlier draft. This article was prepared while the author was on a sabbatical leave from Tel-Aviv University and at the University of North Carolina, Chapel Hill.

Requests for reprints should be sent to Arie W. Kruglanski, who (until August 1976) will be at the Department of Psychology, Vanderbilt University, Nashville, Tennessee 37240.

endogenous-exogenous distinction affords all of the inferences accredited unjustifiably to the internal-external partition between the action's causes. To that end classical attribution experiments that have relied on the internal-external framework are reviewed and reinterpreted, and in some cases novel data are reported that support the endogenous-exogenous distinction. But before a discussion is undertaken of the specific explanatory properties of the endogenous-exogenous framework, it is necessary to consider the more general problem of explanation through attribution. This is attempted in what follows.

Attributional Explanation

The lay person's inference process is assumed to comprise two separate steps, *causal assignment* and *causal explanation*. The first step has been analyzed closely by the attribution theorists (in particular by Kelley, 1967, 1971a, 1971b, 1973). During this stage the attributor aims at identifying the cause of an effect via analysis of variance-like procedures of hypothesis testing (cf. Kelley, 1967). However, little attention has been paid to the second step identified above: the stage of causal explanation. This is the problem of drawing some conclusions about the effect once its cause had been identified. More specifically, attributional explanation is assumed to occur when the effect's determinant has been identified, such that certain inferences follow beyond those contained in the effect notion as such. The attributor begins with an effect, say *b*, and several possible determinants thereof, *X*, *Y*, and *Z*. The causal-assignment stage referred to earlier serves to determine which of these may be most plausibly viewed as *b*'s determinant. Once this has been established, explanation may take place, that is, certain novel inferences may be drawn.

Such inferences may be conveniently classified into two levels, called here the *functional* and the *derivative* levels. On the functional level of explanation (e.g., see Eaker, 1972) the effect (*b*) has been functionally related to some "causal antecedent" (*X*), characterized in empirical or observational (as opposed to theoretical) terms.

For example, in the explanatory statement, "John handed over his wallet because Paul had asked him to do so," Paul's asking would constitute a functional explainer, as it designates an empirical event devoid of conceptual content. The information gain that accrues from deciding that the effect has been functionally related to some empirically defined determinant is confined to the assertion of covariance between the two. Speaking more precisely, everything else being constant, *X* (the determinant) has been identified as the necessary as well as the sufficient condition of *b* (the effect). Thus, the elimination of *X* will do away with *b*, and the reintroduction of *X* will reproduce *b* as well. In terms of our example, all that may be said is that in the specific circumstances Paul's asking (for the wallet) would suffice and be necessary to ensure John's compliance.

A derivative-level explanation presupposes a functional relation, but the effect's determinant is now characterized theoretically (as opposed to empirically on the mere functional level). For instance, John's yielding of the wallet might be explained by reference to the appropriate request coming from an armed bandit. The difference between an empirical and a theoretical characterization of an explainer is that the latter affords derivations that transcend the assertion of covariance between the determinant and the effect, while the former's implications are exhausted with the covariance assertion. In our example, the lay-theoretic concept, *armed-bandit*, affords several informative derivations, for example, that John was experiencing fright while yielding the wallet, that harm might have befallen him had he refused, or that most people in his place would have also complied. To state this differently, on the derivative explanatory level *b*'s possible determinant *X* carries with it a set of implications of which *b* is but one element, say, *b*, *c*, *d*, and *e*, while the determinant *Y* might carry a different set of implications containing *b*, for example, *b*, *e*, *f*, *g*, and *h*. Thus, by identifying *b*'s antecedent as *X* one may derive *c*, *d*, and *e*, whereas by identifying *b*'s antecedent as *Y* the derivations *e*, *f*, *g*, and *h* will follow.

Generally, one may define the degree of explanation as the extent to which inferential gain has been made (or as the size of the implicational set) that accrues from characterizing the antecedent of an effect in a particular way. It follows that explanation at the functional level is weaker than one at the derivative level, and within the latter level the degree of explanation will vary with the number of derivations comprised in (i.e., the theoretical content of) the antecedent term of the explanatory proposition.

The notions above facilitate the analysis of the explanatory properties of the endogenous-exogenous partition. But first let us lay out the rationale for preferring (in some domains of explanation) the latter classification over the current alternative, the distinction between internal and external causes.

Actions Versus Occurrences

It is useful to classify the various effects to which attributional analyses may be applied into two major categories: *occurrences* and *actions*. The fundamental distinction between these two is that actions are voluntary while occurrences are not. In other words, an action is commonly assumed to be determined by the actor's will, while an occurrence is some significant degree independent of the will. For instance, effects such as "running," "stepping on a scale," or "uttering a sound" would qualify as actions, because persons may bring them about simply by deciding to do so. By contrast, "winning a race," "weighing 156 pounds," or "being heard" would represent occurrences, as would "running a fever," "feeling pleased," or "experiencing an acute anxiety." Note that with the former three occurrences some action was presupposed (e.g., winning a race presupposes running it), while none was presupposed with the latter three.

But a common feature of occurrences as a class is that one may not engender them by mere willing: The will to succeed may not guarantee success, nor, alas, may the will to weigh 105 pounds guarantee the appropriate scale reading. Now, it is the occurrences' (relative) independence of the

will that renders them explicable in terms of stable or transient attributes. For example, the occurrence, "P's success on a task," may be explained by this person's attributes of competence or diligence, while the occurrence, "O's experience of fright," may be explained in terms of the situational attribute of darkness. As the foregoing examples illustrate, both persons and environmental entities have attributes; therefore, in explaining occurrences it seems justifiable to distinguish between personal (internal) and environmental (external) causality. Indeed, Weiner, Frieze, Kukla, Reed, Rest, and Rosenbaum (1971) made extensive use of the internal-external distinction in their attributional analyses of success and failure occurrences. But for actions (as opposed to occurrences) the internal-external partition appears less justifiable: All actions are commonly assumed to be determined by the will, and the will is internal to the person; hence, all actions are determined internally. This conclusion serves as a departure point for an attributional theory of actions developed in the remainder of this paper.

ENDOGENOUS-EXOGENOUS DISTINCTION IN NAIVE EXPLANATION OF ACTIONS

In the present section a theory is developed about the inferences that may follow from causally attributing actions in two contrasting ways, designated here as *endogenous* and *exogenous*. The theory will subsequently be applied to diverse attributional findings and will be contrasted with alternative interpretations based on the internal-external framework.

POSTULATE 1. For the naive attributor, actions are explicable in terms of reasons or purposes. That is, the lay explanation of actions is teleological rather than causal.

DERIVATION 1 (from Postulate 1). *Insofar as all actions are (assumed) purposive, any action is either its own end or is a means to attaining some further end.* Now it should be clear that the means-goal categories are conceptually distinct and mutually exclusive: The sense in which something is classified a means is one whereby it is a nongoal. This is not to say that for any ac-

tivity its means versus end status may be clearly ascertained. An activity may be means relative to Goal X and an end relative to Goal Y. For example, "riding a bicycle to school" may constitute a means of transportation, and at the same time an end of physical exercise. Of course, the above example does not imply that the concepts of means and end are nondistinct, any more than the concepts *large* and *small* are in the example of \$1,000 being considered at the same time as large relative to the price of an ice-cream cone and small relative to the price of a mansion.

DEFINITION 1. *Let the term endogenous attribution of an action denote the case wherein an action is attributed to itself as a reason. In other words, an endogenously attributed action is an end in itself.*

DEFINITION 2. *Let the term exogenous attribution of an action denote the case where the action is not an end in itself but rather is a means that mediates a further goal, one exogenous to it.*

POSTULATE 2. *All purposes or intentions are internal to the individual. A goal or an intention is internal in that it presupposes the person. Some intentions may arise from man's biological nature (hunger, thirst), others (achievement, altruism) may arise from internalized social norms.*

DERIVATION 2 (from Postulate 2). *The different goals are not divisible into the internal and external categories. The latter would simply be a null set. It is of interest to note that the use of the internal-external distinction in reference to attributional findings has typically lacked explicit rationale. Rather, it seems to have been grounded in an arbitrary semantic decision to characterize some purposes in a "motive" language (that has an "internal" feel to it), and other purposes in a "goal" or object language (that feels "external").* For instance, in the attributional analyses of findings on task motivation (e.g., Deci, 1971) the actor's interest in the task has been frequently referred to as the internal cause of task performance and the contingent monetary reward as the external cause. But one could readily reverse the labels and call the task

external and the desire for money internal. Obviously, to call a purpose "money" or "desire for money" hardly alters the fact that one is dealing with precisely the same purpose.

Corollary to Derivation 2. *The denial of the internal-external dimension in the case of actions simultaneously denies the possibility of any other variables being systematically related thereto.* Thus, the action's degree of uniqueness (the extent to which it was uniquely undertaken by the actor vs. commonly performed by most people), assumed by the attribution theorists to be a major basis for internal inferences, should not, in fact, relate to internality. Nor may the inferences of intrinsic motivation, subjective freedom, positive attitudes, or sincere reporting relate to internality. In turn, this implies that the uniqueness variable should not be expected to correlate with any of the inferences just listed. Experimental evidence pertinent to this assertion is reviewed in a subsequent section, but now let us turn to the explanatory properties of the endogenous-exogenous distinction.

Functional Level of Explanation (Covariance Level)

DERIVATION 3 (from Definition 1). *An endogenously attributed action will covary with the person's goal but will not covary with the situations.* An action that constitutes an end in itself is completely explicated in terms of this end, provided that this action was situationally possible (e.g., endogenously motivated eating could not take place unless food was situationally present). More specifically, given both the existence of some goal and the situational possibility of its consumption, the endogenously motivated (consummatory) action will be performed regardless of other situational changes that affect neither the presence of the goal nor the possibility of its attainment.

DERIVATION 4 (from Definition 2). *An exogenously attributed action will covary with a specific conjunction of the goal and the situation.* By contrast to an endogenously motivated action, an exogenously motivated one is not fully explicated in terms of its goal (and the situational possibility of

the action). An exogenous action assumes a greater degree of situational determination: The situational constraints implicate a given means (to some end) to be the best available. In Situation A, the means to "having breakfast" may consist of walking to the kitchen; in Situation B, of driving to the hamburger stand, etc.

Derivations 3 and 4 above imply a sense in which exogenously motivated actions are more externally (or situationally) determined than are endogenously motivated ones. This may explain, perhaps, why attribution theorists made the presently denied distinction between internal and external goals (such as "interest in a task" vs. the "contingent payment") where the endogenous-exogenous partition would seem applicable. In any event, while endogenously motivated actions may, in a sense, be also more purely internal than exogenously motivated actions, it is the endogenous-exogenous variable (rather than the internal-external one) that seems crucial to the several inferences from actions that have been demonstrated experimentally. We now turn to the explicit derivation of these inferences.

Derivative Level of Explanation

DERIVATION 5 (from Definition 1). *To the extent that an action is endogenously attributed, it implies the actor's positive affect (enjoyment, contentment, satisfaction, etc.).* This follows from the common definition of the *end* concept as a state of affairs that the actor desires. Its attainment (an endogenously attributed action) represents the fulfillment of one's desire, a situation creditable with positive feelings of enjoyment and satisfaction.

DERIVATION 6 (from Definition 2). *To the extent that an action is exogenously attributed, it implies the actor's negative affect, insofar as means is not something that the actor desires.* In the case of an exogenously motivated action, the actor is doing something that he does not want to do; hence he may not be assumed happy at the time.

Note that the endogenous-exogenous variable is not assumed to constitute the exclusive determinant of satisfaction or positive

affect. An action is endogenous or exogenous with respect to a single goal, while the degree of satisfaction-dissatisfaction is assumed contingent on the totality of the person's goals or needs: notably, on the degree to which they are fulfilled or frustrated. Thus, the present Derivations 5 and 6 need to be supplemented by a *ceteris paribus* proviso, whereby the inferences of satisfaction versus dissatisfaction are a monotonic function of the endogenous-exogenous variable only when the state of the person's remaining goals (needs) is held at a constant level.

POSTULATE 3. *Subjective freedom consists in the attainment of goals.* One does not freely choose to have needs. These are determined for the individual by his biological constitution, societal norms, ethical injunctions, etc. But granting the existence of needs, freedom consists in need fulfillment or goal attainment.

DERIVATION 7 (from Postulate 3 and Definition 1). *To the extent that an action is endogenously attributed, it affords the inference of subjective freedom.*

DERIVATION 8 (from Postulate 3 and Definition 2). *To the extent that an action is exogenously attributed, it affords the inference of compulsion.* In the case of means (vs. end) the actor is not engaging in an action that is inherently desirable: He does not want to, only he is compelled to do it by the existence of his need and by the situational constraints that pose barriers in the way of immediate need satisfaction.

The analysis of inferred freedom contained in Postulate 3 and Derivations 7 and 8 seems closely akin to Steiner's notion of "outcome freedom," which he defined as "high (subjective) probability of obtaining desired outcomes" (1970, p. 189). Obviously, maximal subjective probability of obtaining a desired outcome is experienced upon this outcome's actual attainment, that is, in the case of consummatory or endogenously attributed behaviors. Steiner's "decision freedom" is said to obtain when the individual may believe that "he rather than other people, fate or the press of circumstances" (1970, p. 189) may be credited with his decision. Steiner goes on to illustrate

high (vs. low) decision freedom by the existence of an attractive (vs. an unattractive) alternative to the person's actual action. This strongly suggests that "decision freedom" is also reducible to endogenous attribution. Thus, an unattractive alternative implies an exogenous attribution: The person did not choose the action for its own sake but rather to avoid the unpleasant alternative. Conversely, when the alternative is attractive, endogenous attribution seems indicated, there being no apparent exogenous reason for the action.

DERIVATION 9 (from Definition 1). *In the case of endogenously attributed actions, knowledge of the action's content immediately identifies the action's goal or intention.* For instance, the knowledge that a benefit bestowed on another person was endogenously motivated yields the immediate inference that the actor's underlying intention was benevolence.

DERIVATION 10 (from Definition 2). *In the case of exogenously attributed actions, knowledge of their content does not reveal their underlying intentions.* The specific means chosen (the content of which is known) could, in principle, mediate all possible goals, its choice being situationally determined and not uniquely determined by the underlying goal.

This concludes the present statement of the theory of endogenous attribution. The theory will now be applied to the several areas of empirical research pertinent to the lay explanation of actions.

Evidence for Endogenous Attribution Theory

Numerous experiments on lay inferences from actions have been interpreted in terms of the internal-external distinction mentioned earlier. Typically, the reasons for designating specific experimental events as the internal or the external causes of actions have not been explicitly given. However, one of possible two criteria seems implicit in all such classifications. The first is based on the idea that the lay attributor intuitively classifies certain reasons for action (e.g., monetary payment) as generically external to the actor and other reasons (e.g., interest

in a task) as generically internal. The second criterion assumes that the attributor systematically applies the uniqueness standard to arrive at internality designations: That cause is considered internal whose effect is unique to the actor versus being common to many persons (Kelley, 1967).

The present theory of endogenous attribution suggests that in explaining actions the layman draws the endogenous-exogenous distinction between the actions' causes (reasons) rather than the internal-external distinction. In the following review of attributional research on actions, the case for the present theoretical scheme will rest on both conceptual and empirical considerations that favor the former over the latter partition. On the conceptual level, two difficulties with current applications of the internal-external framework will be identified: the problems of (a) criterion and (b) inference. The problem of criterion is that the two ways of rendering internality-externality attributions implied earlier (the ideas of generically internal and external causes and of the uniqueness standard) do not appear to coincide, as might be hoped for, but rather seem to be mutually independent. Thus, any generically internal cause could be implicated as external on grounds of nonuniqueness, and any generically external cause as internal. By contrast, the conceptual criterion for the endogenous-exogenous distinction is relatively straightforward, as it rests on the common differentiation between means and ends.

The problem of inference is that in typical applications of the internal-external framework, it is neither explicitly elaborated nor is it otherwise plain how the attributions of internality (externality) may imply the various inferences with which they have been credited. By contrast, the present Derivations 3-10 specifically deduce the various inferences afforded by the endogenous-exogenous classification.

The main empirical considerations in support of the present theory include the arguments that (a) in all those experiments where the subjects' inferences were coordinated to the internal-external distinction a reinterpretation is possible congruent with

the theory of endogenous attribution; (b) specific experimental attempts to vary the internal-external factor (either by manipulating the uniqueness variable or the generically internal or external causes) failed to yield the inferences claimed to be implied by this partition; and (c) experiments that specifically varied the endogenous-exogenous factor yielded results that unexceptionally corroborated the present theory.

Let us now directly turn to attributional research on the lay explanation of actions. The specific topics to be reviewed include (a) intrinsic motivation, (b) subjective freedom, (c) benevolent and malevolent intentions, (d) inferred sincerity, and (e) escape behavior.

Intrinsic motivation. Several recent experiments have demonstrated a negative relation between the existence of contingent, frequently monetary, reward and the subject's degree of intrinsic motivation for the task (e.g., Deci, 1971, 1972a, 1972b; Greene & Lepper, 1974; Kruglanski, Alon, & Lewis, 1972; Kruglanski, Friedman, & Zeevi, 1971; Kruglanski, Riter, Amitai, Margolin, Shabtai, & Zaksh, 1975; Lepper, Green, & Nisbett, 1972). These findings have been sometimes interpreted to mean that the actor infers his own intrinsic motivation when he attributes his performance of the task internally (cf. Deci, 1974), for example, to his interest in the task. The inference of intrinsic motivation was assumed less likely in the reward versus the no-reward condition, because the former case but not the latter also contains plausible external reasons for task performance, that is, the contingent rewards. Such analyses as the above have usually failed to clarify the criterion for designating the contingent reward (e.g., the money) as the external cause of the action. Rather, in these cases the cause's generic nature was somehow intuited to belong in the external category. Alternatively, the widespread appeal of money could justify its alleged externality on grounds of its nonuniqueness as a motivator. It should be recognized, however, that the two criteria just implied are orthogonal and may not be expected to manifest unexceptional correspondence. For instance, a person could be

uniquely greedy and in return for money undertake a task rejected by everyone else. From the uniqueness perspective this would establish money as an internal cause of this person's action, that is, contravert the suggestion that money is generically external.

Furthermore, the various analyses of intrinsic motivation in terms of the internal attribution fail to cogently rationalize the implied relation between the two. Yet, counterexamples easily come to mind. For instance, our greedy individual above would seem to be internally motivated (on uniqueness grounds), yet would seem quite unlikely to be credited with a great deal of intrinsic motivation toward the task that he performed for money. The generic standard for internality-externality is similarly susceptible to counterexamples with respect to the inferences of intrinsic motivation. On this criterion money would represent a typically external cause, so that behaviors prompted thereby should lead to the inferences of extrinsic motivation expressed, for example, in disenjoyment or dissatisfaction. But acts to which the monetary rewards are inherent, such as playing poker, performing a business transaction, or receiving one's paycheck, are certainly attributable to monetary causes without implying disenjoyment.

The last examples naturally lead to the present conceptual framework, whereby the inference of intrinsic motivation follows from an endogenous attribution (Derivation 5). Whenever the actor's performance of his task appears endogenously motivated (i.e., task performance appears to be an end in itself) the implication of enjoyment, interest, etc. derives from the notion of goal attainment (consummatory behavior). By contrast, when there exists a plausible exogenous reason for task performance (the actor's reason is not inherent to the task but is made arbitrarily contingent on task performance), the attributor will partially discount task-endogenous reasons; hence the inference of enjoyment, interest, etc. will be less likely to follow. Thus, the present analysis suggests that the inference of extrinsic motivation in the reward condition does not follow from (a) the fact that money was the goal of one's activity nor (b) the fact that

TABLE 1
AVERAGE DEGREES OF INTRINSIC MOTIVATION
(EXPERIMENT 1)

Condition	Money endogenous	Money exogenous
Payment present	5.44	4.66
Payment absent	2.25	5.60

Note. Higher figures denote a greater degree of intrinsic motivation on a 7-point scale. From "Can Money Enhance Intrinsic Motivation? A Test of the Content-Consequence Hypothesis" by A. W. Kruglanski, A. Riter, A. Amitai, B. Margolin, L. Shabtai, and D. Zaksh, *Journal of Personality and Social Psychology*, 1975, 31, 744-750. Copyright 1975 by the American Psychological Association. Reprinted by permission.

the goal was nonunique; but rather that the activity was exogenously attributed, that is, was a means, rather than an end in itself.

Recently, this author has completed two experimental sequences of direct relevance to the discussion above. The results of this research suggest that (a) money, a cause frequently viewed as external to the person, might be endogenous to an activity, in which case its presence would enhance rather than suppress intrinsic motivation; and (b) the degree of intrinsic motivation does not seem to be a function of the uniqueness variable. Regardless of the degree of uniqueness, that of intrinsic motivation is greater when the effect (action) is attributable endogenously rather than exogenously. Let us now consider at some length the experimental evidence for these propositions.

In the first experimental paradigm (Kruglanski et al., 1975) monetary rewards were in the one case inherent to the task content (money-endogenous condition), while in the second case they were unrelated to the task (money-exogenous condition). Within each of these conditions half of the subjects performed the task in return for monetary rewards and the remaining half for no rewards. Specifically, in the money-endogenous condition of Experiment 1 the subjects played with the experimenter a coin-toss guessing game (heads or tails). As this game is usually played for actual money, it had been assumed (and confirmed on subsequent checks) that money would be apprehended as a game-endogenous feature. In the money-exogenous condition the subjects played with the experimenter a game of

model construction, an activity hardly ever associated with money. In each of the task conditions half of the subjects received monetary rewards for correct performance (the payment-present conditions), and the remaining half received points (payment-absent conditions). The degree of intrinsic motivation was inferred from the appropriate index based on the subjects' reported interest in the game and preference for it over possible alternatives. The pertinent data are displayed in Table 1. As shown, when money is endogenous to the task its presence enhances intrinsic motivation ($p < .001$), whereas when it is exogenous to the task its presence depresses intrinsic motivation ($p < .05$)—this in contrast to the implication that intrinsic motivation is always lowered when an external cause of actions (such as money) appears in evidence.

The confidence in the above results is augmented by the findings of Experiment 2, where the money-endogenous and money-exogenous tasks were less dissimilar than their counterparts in Experiment 1. Specifically, in both game conditions of Experiment 2 the subjects' actual activities were identical, only they were interpreted differently: as stock market operations in the money-endogenous conditions and as athletics management in the money-exogenous condition. As in Experiment 1, the sub-

TABLE 2
GAME PREFERENCES AND RATED INTEREST
(EXPERIMENT 2)

Measure	Money endogenous	Money exogenous
Proportions of subjects preferring the same (vs. another) game		
Payment present	.65	.35
Payment absent	.25	.70
Average rated interest		
Payment present	5.75	4.43
Payment absent	4.31	5.30

Note. Higher figures denote a higher rated interest on a 7-point scale. From "Can Money Enhance Intrinsic Motivation? A Test of the Content-Consequence Hypothesis" by A. W. Kruglanski, A. Riter, A. Amitai, B. Margolin, L. Shabtai, and D. Zaksh, *Journal of Personality and Social Psychology*, 1975, 31, 744-750. Copyright 1975 by the American Psychological Association. Reprinted by permission.

jects' successful performance (accumulation of stock or athletic credits) was rewarded by money in the payment-present conditions and by points in the payment-absent conditions. Intrinsic motivation was inferred from (a) the proportions of subjects who preferred to continue with the same game versus trying an unknown alternative, and (b) interest and enjoyment ratings. The relevant data are shown in Table 2. As expected, a significantly greater proportion of the subjects preferred the money-endogenous game in the payment-present condition ($p < .01$) and preferred the money-exogenous game in the payment-absent condition ($p < .02$). Similarly, the money-endogenous game was rated as significantly more interesting in the payment-present condition ($p < .01$), whereas the money-exogenous game was rated as more interesting in the money-absent condition ($p < .05$). In sum then, the findings of both experiments are consistent with the present theory of endogenous attribution (specifically with Derivation 5) and inconsistent with the idea that intrinsic motivation follows from internal attribution and is lowered by the presence of a generically external cause such as money.

The second series of studies (Kruglanski, Note 1) examined the uniqueness standard for internality and the endogenous-exogenous variable as they may affect the inferences of intrinsic motivation and freedom. In the first experiment the subjects were furnished with information regarding the target person's uniqueness on some characteristic. In one half of the cases the target person was presented as unique on that characteristic and in the remaining half of the cases as nonunique. The target person was then portrayed to have acted in some manner, such that in one half of the cases the individual's unique or nonunique characteristic suggested an endogenous attribution, while in the remaining half of the cases it suggested an exogenous attribution.

Thus, in one of the four hypothetical situations presented to the subjects, the actor, a high school student, decides to attend a Sunday concert performed by an out-of-town classical ensemble. In the endogenous condition the actor is known for his delight in

TABLE 3
ENJOYMENT AND FREEDOM RATINGS
(EXPERIMENT 1)

Measure	Condition	
	Endogenous	Exogenous
Inferred enjoyment		
Internal condition	6.38	4.00
External condition	6.31	3.87
Inferred freedom		
Internal condition	6.57	3.88
External condition	6.15	3.88

Note. The higher the figures, the greater the inferred enjoyment or freedom. From Kruglanski (Note 1).

classical music, which is either unique to him or is common among his fellow students. In the exogenous conditions, Ann, an attractive new girl, is also known to attend the concert, and the protagonist has an interest in Ann that is either unique to him or is commonly shared among the boys at his school. The subjects responded to four hypothetical situations in accordance with a balanced Latin square design, where each situation represented one of the four possible combinations of the endogeneity and the uniqueness variables. After reading each text, the subjects answered a brief questionnaire that included checks on the uniqueness manipulation and questions about the actor's perceived enjoyment of the action (e.g., listening to the concert) and his subjective freedom while pursuing it. It will be noted that the uniqueness manipulation was highly effective: The subjects rated the protagonist's actions as more distinctive of his person in the unique (vs. the nonunique) condition ($p < .001$). The enjoyment data (collapsed across the four situations) are displayed in the upper row of Table 3. As shown, greater enjoyment of the activities was reported in the endogenous (vs. the exogenous) conditions ($p < .001$), while the uniqueness manipulation yielded no effects with respect to task enjoyment.

The possible argument could be raised that the uniqueness variable did not affect the results of Experiment 1 because the actor's reason for his behavior was clearly identified. Perhaps uniqueness may signify

TABLE 4
ENJOYMENT AND FREEDOM RATINGS
(EXPERIMENT 2)

Measure	Condition			
	Internal	External	Endogenous	Exogenous
Inferred enjoyment	5.37	5.00	6.45	3.87
Inferred freedom	5.54	5.50	6.36	4.29

Note. The higher the ratings, the greater the inferred enjoyment or freedom. From Kruglanski (Note 1).

internality only when the person's specific reason for his action is not otherwise implied. The foregoing possibility was explored in Experiment 2 of the sequence. Subjects responded to hypothetical situations in which the actor had two possible reasons for a given behavior, one endogenous to this behavior and the other exogenous. Additional information was provided that suggested the actor's reason to be either (a) the endogenous one, (b) the exogenous one, (c) unique to the actor, or (d) nonunique to him. (Note that in the latter two cases the sole information provided was about the uniqueness of the action without the specific reason being identified in any way.)

Thus, in one of the hypothetical situations an employee is having a dinner with his boss and decides to order a delicious pie for dessert (endogenous reason for the order), exactly as ordered by the boss a moment ago (exogenous reason for the order). Additional information furnished by a professional psychologist suggests this person's action to be (a) highly unique to his personality, (b) nonunique, as most people in this situation would have done the same, (c) prompted endogenously (i.e., solely by the pie's apparent properties), (d) prompted exogenously (i.e., solely by the boss's preferences). Based on a balanced, Latin square design, subjects responded to four situations, each representing one of four experimental conditions (e.g., the unique, nonunique, endogenous, and exogenous conditions). The dependent variables were, again, the ratings of the actor's enjoyment and of his subjective freedom with respect to the action. The enjoyment data are

summarized in the first row of Table 4. Again, the endogenous-exogenous variable strongly affected the results but not the uniqueness variable. In particular, the actor's inferred enjoyment was (a) higher in the endogenous case than in either the internal or the external cases ($p < .01$), (b) lower in the exogenous case than in either the internal or the external cases ($p < .05$), and (c) no different in the unique and the nonunique conditions.

In Experiment 3 of this series subjects assessed the information value for enjoyment and freedom inferences of questions posing the choice between (a) endogenous versus exogenous, and (b) uniquely personal versus uniquely environmental origin of the actor's reason for his behavior. The latter choice was modelled after the example of the internal-external dilemma given sometimes by the attribution theorists. For instance, Kelley (1967, p. 194) describes the movie viewer who wonders whether his enjoyment has been produced by his own distinctive property ("a specific kind of desire relevant to movies") or by the movie's property ("that it is intrinsically enjoyable"). In the present case, one of the four hypothetical situations used in Experiment 3 portrayed a protagonist, X, who decided to attend movie Y. In the personal versus environmental version the question whose information value the subjects were to assess was whether X's reason had to do with his uniquely personal preferences or with the movie's properties. In the endogenous-

TABLE 5
RATED INFORMATION VALUE OF INTERNAL-EXTERNAL VERSUS ENDOGENOUS-EXOGENOUS CHOICES FOR ENJOYMENT AND FREEDOM INFERENCES
(EXPERIMENT 3)

Measure	Condition	
	Internal-external	Endogenous-exogenous
Enjoyment inference	3.62	5.12
Freedom inference	3.54	5.41

Note. The higher the figures, the greater the perceived information value. From Kruglanski (Note 1).

exogenous version the question to be assessed was of whether X's reason for attendance had to do with the movie's properties or with the fact that his boss had invited him to come along for this movie. Each of the three hypothetical situations was presented to the subjects twice in a counterbalanced order: once with the personal-environmental question and once with the endogenous-exogenous one. The pertinent results (collapsed across the three situations) are displayed in the upper row of Table 5. As shown, the subjects rated the endogenous-exogenous question as significantly more informative for enjoyment inferences ($p < .01$) than they did the personal-environmental question.

Finally, relevant to the present concern is the "forced-compliance" study by Cooper, Jones, and Tuller (1972), where uniqueness versus nonuniqueness of a person's compliance to deliver a counterattitudinal essay on some issue was orthogonally crossed with the magnitude of exogenous reward promised him for doing so. It turned out that a more positive attitude on this issue was expressed by subjects in the case of a small exogenous reward (versus a large one), that is, where endogenous attributions would seem more likely. By contrast, the act's degree of uniqueness had no appreciable effects on attitudes.

Subjective freedom. It has been argued that freedom or volition of the actor may be inferred from an internal attribution, that is, to the extent that the act may be attributed to the person uniquely. For example, Kelley (1967, p. 218), suggested that "you can be said to exercise choice only if you do something *different* from what everyone else would do in the same situation." But it would seem that some nonunique choices may appear (and be experienced) as quite free, and some unique ones as unfree. For instance, suppose that a millionaire, notorious for his eccentricity (say Howard Hughes) approached you one day with the free gift offer of \$1,000,000. Most of us would probably hasten to accept the gift without feeling unfree in the least. On the other hand, a person's unique fear of authority might prompt him to uniquely abide

by unreasonable regulations, unheeded by anyone else, and experience a strong sense of compulsion while doing so.

On the level of empirical evidence, Kruglanski and Cohen (1973) claimed specific support for the uniquely personal rendition of perceived freedom. But a closer scrutiny of these authors' procedures suggests that their findings may better be explained by reference to the endogenous-exogenous distinction. Kruglanski and Cohen (1973) operationalized person causation by the experimental condition where the target person's act (choice of an essay topic) was consistent with his unique character versus being inconsistent with his character in the external-causation case. Subjects attributed significantly greater freedom to the target person in the in-character versus the out-of-character condition.

But rather than viewing the in-character condition as one that operationally defines uniquely-personal causation, and the out-of-character condition as one that defines the absence of unique causation, it is possible to interpret the former as implying an endogenous attribution and the latter an exogenous one. Thus, in the in-character condition, knowledge of the protagonist's preferences suggests that the act produced (choice of an essay) was endogenously motivated, whereas in the out-of-character condition the same prior knowledge suggests that the act may not have been its own end; hence some aspect exogenous to this act must have been the reason for its performance. Thus, in the Kruglanski and Cohen experiment, uniquely-personal causation was confounded with endogenous attribution, and either conceptualization could explain the findings. But in the experimental series by Kruglanski (Note 1) described above, freedom inferences (as well as those of task enjoyment) were significantly ($p < .01$) affected by the endogenous-exogenous distinction but not at all affected by the uniqueness variable (Experiments 1 and 2). Furthermore, questions posing the endogenous-exogenous choice were rated as significantly ($p < .01$) more informative for freedom inferences than questions posing the internal-external choice. The data relevant to the above assertions

are displayed in the lower rows of Tables 3, 4, and 5. Finally, it is of interest to note that within Experiments 1 and 2, the ratings of freedom and enjoyment exhibited a high degree of positive correlation, as did the informativeness ratings for freedom and enjoyment inferences in Experiment 3 ($p < .01$ in all cases), a finding consistent with the idea that freedom and enjoyment inferences follow from the same (presumably) endogenous attribution.

To conclude, then, both previous results and novel experimental findings support Derivations 4 and 5 of the present theory, whereby the inferences of intrinsic motivation and of freedom follow from the endogenous attribution of actions. Similarly consistent with the present theory are the recurrent failures to corroborate the contention that internal attribution (e.g., as operationalized in terms of the uniqueness criterion) mediates the freedom or the intrinsic motivation inferences.

Benevolent and malevolent intentions. The rendition of intentions as benevolent or malevolent has been analyzed by reference to the internal-external distinction. In a typical experiment, the actor commits a helpful or a harmful deed (a) under a condition where an external reason is apparent or (b) in the absence of an external reason. External reasons have included the status of the act's recipient (Thibaut & Riecken, 1955), the experimenter's instruction to perform the act (cf. Goranson & Berkowitz, 1966), "a widely recognized physical cause; a compelling institutional arrangement or rule; a factor in the frustrator's psychological makeup, such as the mental illness of a person who attacks you . . . a lack of ability on the part of a person upon whose success one is dependent . . . an inadequacy in the prior behavior of the person subjected to the frustration" (Kelley, 1971a, p. 15). It turns out (see Kelley, 1971a, for a review) that in the absence of such external reasons, a helpful or harmful act elicits a greater degree of reciprocation from the target person. This finding has been interpreted to mean that internally attributed benevolent or malevolent acts implicate be-

nevolence or malevolence as their underlying intention.

The diversity of external reasons illustrated above raises again the question of criterion: What common aspect of these different justifications may account for their externality (or the internality of others). Little specific discussion of this problem exists in the literature, but it is possible to apply here our two earlier interpretations, whereby the attributor may apply either the intuitive criterion of generic internality (externality) or the more systematic standard of uniqueness. It is worthwhile to reiterate that the two interpretations need not coincide. For instance, a person could be particularly sensitive to social status, and hence go to unique pains to please a high-status figure. This would suggest internal causation on grounds of uniqueness, even though the cause in question (status) has been customarily referred to as external (cf. Thibaut & Riecken, 1955).

The above example also highlights the problem of inference with respect to the internal-external partition: Internal attribution has been assumed to mediate the reciprocation of one's helpful or harmful acts, yet it seems unlikely that an actor's unique enactment of a benevolent deed should result in much reciprocation, for example, when that deed is assignable to the person's unusual craving for status.

In the empirical domain, inexplicable in terms of the uniqueness principle is the persistent finding that less reciprocation follows when the actor's helping behavior seems inappropriate, hence suggesting "that the actor may be peculiar, stupid, or not completely rational" (Kelley, 1971a, p. 16). Clearly, inappropriate or irrational behaviors (whether benevolent or malevolent) are ipso facto unique, and hence should increase the probability of internal attribution, yet low reciprocation resulted. Kelley viewed such findings as "exception[s] to the rule that helpful or affiliative behavior attributed to the person will elicit more liking than if attributed to the situation" (1971a, p. 15). However, within the present endogenous-exogenous interpretation such results need not be

viewed as exceptional. Let us, therefore, consider the findings on reciprocated harm or benefit in light of the latter distinction.

It is presently proposed that the presence of justifying reasons represents the case wherein the act (the benefit or the harm) may not be readily assigned endogenously, but rather may be perceived as a means toward the attainment of an exogenous goal (the justifying reason). Now, when the act's content is known to be benevolent or malevolent, an endogenous attribution implicates the actor's intention as benevolence or malevolence (Derivations 9 and 10). It is the latter inference of intention that may elicit reciprocation from the act's target. The above analysis suggests, for example, that the benevolence inferences and reciprocated benevolence (assumed to follow from an endogenous attribution of a benevolent action) ought to be correlated with the inferences of the actor's freedom (assumed to follow from an endogenous attribution of any action). Indeed, in at least one experiment such correspondence was reported. Specifically, Thibaut and Riecken (1955) found that when a helpful act was performed in the absence (vs. the presence) of an exogenous reason (approval by a high status figure) the actor was (a) credited with greater freedom (was perceived as wanting to rather than being coerced to help) and also was (b) liked more (reciprocated benevolence).

As noted earlier, the uniqueness criterion for internality confronts a special difficulty in the case of irrational behaviors (or those associated with a psychological defect in the actor). Such behaviors are highly unique (which implies internality) yet have typically resulted in low reciprocation—reserved for externality. Therefore, of particular interest is the possibility that the theory of endogenous attribution may handle the case of irrational or pathological acts. Specifically, the irrational actor could misperceive the situation and include in his evaluation a variety of exogenous justifications evidently absent for a more rational perceiver (e.g., the paranoiac may inappropriately perceive the other person as boding him ill; such an

evaluation could then provide a possible exogenous reason to which his aggressive acts could be attributed). Similarly, when performance of a given act seems clearly inappropriate, the actor could be suspected of misperceiving the meaning of this act. This would render inapplicable Derivations 9 and 10 of the present theory, thus delimiting the implications of endogenous attribution to the inferences of freedom and motivation to perform the act as such. But in the absence of clarity that the actor understood the act as benevolent or malevolent, an inference that the actor's intention was to bestow benefit or harm is not unambiguously indicated.

To sum up, then, it has been argued that the data on reciprocated harm and benefit are accountable for within the endogenous-exogenous framework and not readily explicable in terms of the uniqueness variable or the internal-external partition as such. An endogenously motivated favor or harm implicates not only the actor's volition (as in the Thibaut & Riecken, 1955, study) but also his intention as benevolence or malevolence, inferences which, in turn, may mediate reciprocal behavior from the act's target.

Inferred sincerity. In a widely known argument, Daryl Bem (1965, 1967) accounted in attributional terms for the negative relation between the magnitude of inducement for making a counterattitudinal statement and subsequent attitude change toward the statement's content (cf. Festinger & Carlsmith, 1959). To that end Bem employed Skinner's (1957) distinction between *tact* and *mand*, where *tact* has been defined as "a verbal response . . . under the discriminative control of some portion of the environment" and *mand* "as a verbal response under the control of a specific reinforcing contingency" (Bem, 1965, p. 200). When the magnitude of inducement is high, the reward contingency is clear, hence a *mand* inference should be made, which implies that the statement may not have reflected the (subjectively) true properties of the entity about which it has spoken. By contrast, with a low inducement a *tact* in-

ference is more likely, and it should afford the inference of a sincere report.

Note that the tact-mand distinction differentiates between portions of the (external) environment and assumes a causal explanatory scheme. In other words, a person's utterances are assumed to be caused by various parts of the environment or to constitute responses to environmental stimuli. This contrasts with the teleological account of actions (including utterances) proposed herein. But from a different vantage, there exists an important similarity between the tact-mand and the endogenous-exogenous distinctions. Tacting refers to utterances prompted by endogenous aspects of the statement's object, while manding implies utterances prompted by goals (e.g., reward contingencies) exogenous to the statement's object. Now in the same way that endogenously motivated benefit or harm implies the actor's end to be benevolence or malevolence, so endogenously motivated reports may implicate sincerity ("telling it like it is") as their underlying goal (Derivations 9 and 10). In other words, a report is a description of some reality, and an endogenously motivated report implies that the actor's sole objective is to describe a reality, that is, report on what to him appears to be the case.

Note that a quite different theme looms prominent in Kelley's (1967) interpretation of the relation between inducement and attitude change. Kelley viewed as particularly significant the fact that "strong reinforcement contingencies suggest the notion that most or all persons in the same situation would make a response" (1967, p. 217). He also emphasized the idea that

the judgment that the attitude of a [low inducement] complier is more favorable than that of a [high inducement] complier is probably associated with assumptions that there is a distribution of opinion toward the task, and only the more favorable subjects complied in the [low inducement] case and almost all, favorable or not, complied in the [high inducement] case. (p. 226)

The foregoing analysis has two implications:

1. The distinction proposed therein is not between entity versus extrinsic reward (tact

vs. mand), but rather between internally versus externally assigned causality as defined from the uniqueness perspective. The low-inducement case is assumed to suggest an internal attribution, whereas the high-inducement case is assumed to suggest an external attribution (as everyone in the same situation is assumed to comply).

2. Within the above analysis, internal or unique attribution of an utterance is assumed to reflect the speaker's sincerity or favorability to his pronouncement.

But the proposed linkage between internal attribution and the inferences of sincerity of favorability has not been furnished with an explicit rationale, nor does it appear to be intuitively evident. On the contrary, the consensus criterion of attribution validity (Kelley, 1967) suggests that unique pronouncements (those that deviate from the consensus) would be considered false, thus (possibly) insincere. Similarly, unique (hence, internally based) compliance need not reflect the actor's favorability toward his deed but could, alternatively, signify his special craving for the contingent reward, or his unique fear of the contingent punishment. Indeed, the above-mentioned research by Cooper et al. (1972), and by Kruglanski (Note 1) has failed to detect a relation between the uniqueness of an act and the actor's inferred favorability thereto.

As another example of the subjects' alleged reliance on the internal-external scheme in arriving at sincerity judgments, consider the experiment by Jones, Davis, and Gergen (1961). In this study, a fictitious interviewee for a job described himself in a manner that was either consistent or inconsistent with the job requirements. It turned out that subjects were more prone to believe the interviewee's statements in the latter versus the former case. In interpreting this finding, Kelley suggested that the contributor's behavior was prompted by an internal cause or by an external one—the demand characteristics of the job interview situation. Kelley further argued that in those conditions where the job applicant had portrayed himself inconsistently with the job requirements "the behavior should be attrib-

uted to the person himself and from it his properties should be inferred" (1971a, p. 9). But such an interpretation suffers from the by now familiar difficulties of the internal-external partition. First, the classification of role demands as external to the person lacks explicit rationale. For instance, compliance with role demands could be unique (hence, internally based) because of some actor's unique craving for role-associated rewards. Second, the suggestion that internal (or unique) attribution leads to the inferences of sincerity is similarly unexplained. As noted earlier, the uniqueness of an actor's statements (i.e., his deviation from consensus) could well intimate a deceptive intent on his part, rather than sincerity.

A quite simple account of the Jones et al. (1961) results is possible in terms of the present theory of endogenous attribution: Behaviors inconsistent with the job requirements are more readily attributable endogenously than behaviors consistent with the job requirements, which suggest the additional possibility of an ulterior (exogenous) purpose. Furthermore, characterization of the behavior in question as a self-report (i.e., a description of some reality about oneself) allows, in the case of endogenous attribution, identification of the speaker's specific intentions as that of describing oneself sincerely (Derivations 10 and 11).

It is of more than curious interest that the situation concocted by Jones et al. (1961) easily lends itself to the mistaken uniqueness interpretation for two incidental reasons: (a) The interviewee is making a statement about himself; in other words, endogenous aspects of the statement's content concern his unique person; and (b) the exogenous reward (getting the job) is something that appears to be of high general desirability (i.e., suggesting a nonunique and hence external attribution). Thus, within this situation endogenous attribution could be also characterized as attribution to the unique person, as "person" constitutes here an endogenous aspect of the statement's object (simply, the statement is about the person). But it is important to realize that the sin-

cerity inference follows from attribution to the person as endogenous aspect of the statement's object, not to the person of the speaker as such. Similarly, in this instance the exogenous reward (lowering the confidence in sincerity) is confounded with an implied external attribution. However, the crucial factor is that the attribution is exogenous not that it is external (or non-unique).

To recapitulate, the theoretical implication suggested by the preceding discussion is that verbal descriptive statements attributed endogenously are interpreted as sincere or truthful. To the extent that an exogenous element of the situation (e.g., a contingent reward or threat) appears to constitute the statement's end, the statement will be interpreted as less sincere. As in our discussion of acts classified as beneficial or harmful for the recipient, so in the case of reports; endogenous attribution implies more than the actor's freedom and satisfaction with the behavior (see Derivations 9 and 10). In this instance, characterization of the behavior as a report (or a description of a reality) allows one to infer the endogenous reason for such behavior, namely, one of reporting what, to the speaker, appears to be the case. It is precisely this latter assumption that may allow the interpretation of endogenously prompted statements as sincere.

Escape behavior. In an experiment by Bandler, Madaras, and Bem (1968), subjects in one condition were led to believe that their escape of electric shock occurred under free choice, while subjects in another condition were induced to think that their escape was mandatory. Subsequently, individuals in the first category (free choice) rated the shock as more painful than did their counterparts in the second category. Nisbett and Valins (1971) accounted for these data in terms of the internal-external division, arguing that in the choice condition, "it was expected that the subjects would interpret their escape behavior as self-produced: subjects should infer from this behavior that the shocks were indeed rather painful" (p. 2). By contrast, in the no-choice condition, escape behavior is pre-

sumed to be under external (or environmental) control (p. 3). But again the conceptual difficulties with the internal-external distinction are vivid: (a) The unclarity of the criterion for classifying the experimenter's demands (rather than, say, the shock) as external; (b) the possibility of the experimental demands providing a unique (hence, internal) motivation for a uniquely demand-sensitive actor; and (c) the lack of apparent rationale for deducing the inference of painfulness from the attribution of internality. Thus, our demand-sensitive individual could uniquely (hence, internally) escape a shock because of someone's request (unheeded by anyone else), yet no inference of pain is likely to be made in such an event.

Within the present conceptual framework, escape in the choice condition can be coordinated with the notion of endogenously motivated activity: Cessation of the shock is an end in itself. By contrast, in the no-choice condition escape is less clearly its own *raison d'être*, as there also exists the (exogenous) experimental instruction to escape. Furthermore (as Derivations 9 and 10 suggest), knowledge that the act's content was cessation of shock, allows, in the case of endogenous attribution, the inference that the underlying intention was to terminate the pain. Hence greater painfulness of shock is inferred in the choice than in the no-choice condition, where the act's intention is more difficult to identify.

Conditions of Endogenous Attribution

The foregoing discussion suggests that the endogenous-exogenous partition may mediate such significant phenomena as intrinsic motivation, subjective freedom, and the perception of intention. Thus, it is particularly important to identify the conditions in which the endogenous (or exogenous) attributions will be made. The purpose of the following sections is to adumbrate possible classes of such conditions. This analysis is anchored in Definitions 1 and 2 given earlier, whereby an endogenous attribution is made when the action is considered to be its own reason, and an exogenous attribution is made when

the reason is foreign (exogenous) to the action. The above definitions contain three distinct components: the action, its cause (or reason), and the relation (of endogeneity or exogeneity) assumed to connect the two. Correspondingly, the conditions of endogenous (exogenous) attribution are naturally classifiable into those affecting the action's perceived identity, the action's perceived causality, and the reason's perceived endogeneity to the action. In what follows, the causality category is considered first. This class of conditions may already claim some empirical support, whereas the remaining two categories (identity and endogeneity) constitute at present novel theoretical proposals intended as heuristics to further inquiry.

The causality question. The problem of causal assignment is a general one within the attribution process and is not unique to the endogenous-exogenous distinction. This is the problem of choosing among several causal hypotheses advanced to explain some effect. The attributor's selection of a given hypothesis may be regulated by several principles identified by Kelley (1967, 1971a), in particular the discounting principle and that of corroboration via attributional criteria. According to the discounting principle, the attributor's confidence in any causal hypothesis is inversely related to the number of hypotheses entertained. Thus, confidence in an endogenous explanation will be higher in the absence (vs. the presence) of additional, exogenous explanations of an action. Indeed, all of the findings reviewed earlier may be interpreted to illustrate the workings of the discounting principle. In the basic experimental paradigm the actor in one condition appears to have only an endogenous reason for his behavior, while in another condition he seems to have an exogenous reason as well. The recurrent finding that inferences deducible from the notion of endogenous attribution are stronger in the former (vs. the latter) case support, therefore, the discounting principle of causal assignment.

The discussion thus far pertains to variation in the number of plausible hypotheses.

By contrast, given a constant number of hypotheses, their relative validity may be assessed by reference to attributional criteria of corroborating evidence. More specifically, corroborating evidence for a given causal hypothesis could assume, for example, the forms of consensus information, consistency over occasions or consistency over modality (Kelley, 1967). Alternatively, corroboration of a given hypothesis could be achieved via the convergence of several independent indicators (see, e.g., Ajzen & Fishbein, 1975; Campbell, 1963). For example, consensus information was manipulated by Kiesler, Nisbett, and Zanna (1969) in such a way that pronouncements by a confederate regarding the reason for his behavior resulted in similar attributions by the subject. Consistency (over occasions) figured in the experiment by Kruglanski and Cohen (1973) reviewed earlier, where knowledge of the actor's past preferences implicated a given reason as a plausible explanation of his current action. Finally, the convergence of information could have been involved in the study by Nisbett and Schachter (1966). These investigators found that the judged intensity of shock-produced pain can be manipulated by providing the individual with an alternative explanation for the effect (physiological arousal) he was experiencing. As it turned out, this effect obtained only in the low-fear condition of the Nisbett and Schachter experiment and not in the high-fear condition. According to the present interpretation, the high-fear condition may represent the case where several bodily components of fear (e.g., the sinking feeling in one's stomach, the sweating of one's palms) may converge on the hypothesis that shock was responsible for the arousal; hence, attribution to shock might override possible alternative explanations manipulated by the experimenters.

To summarize, then, the present discussion of the causality problem suggests that endogenous attribution is likely to be rendered when (a) no exogenous reasons are likely to be entertained by the attributor, and (b) when there is evidence (e.g., of the consensus, consistency, or convergence kind)

that corroborates the endogenous explanation and/or falsifies exogenous explanations of the action.

The identity question. The problem of causality, above, has received ample attention within attribution theory. By contrast, little explicit discussion has been devoted to the problem of identity, that is, to the way an effect (to be causally explained) is conceptualized. It is presently suggested that the lay person's knowledge of the effect's identity is hypothetical in the same sense that his knowledge of causality is. In other words, an individual confronted with an event or a sequence of events is assumed to formulate hypotheses about the conceptual identity (meaning) of these phenomena. The foregoing implies that the same physical events are amenable to several disparate identifications. Pertinent to the present concern, the same action may be conceptualized in a variety of different ways, and the specific conceptualization that the attributor decides to adapt will influence the likelihood of the action being assigned endogenously or exogenously.

For example, the same sequence of behaviors might be variously interpreted to mean that the actor is (a) sweeping the floor, (b) practicing his shuffleboard game, (c) helping his mother, or (d) earning money. Obviously, when the action is conceptualized in one way an endogenous attribution thereof needs to assume a quite different purpose on the part of the actor. Thus, if the actor's behavior is conceptualized as helping his mother and its purpose as earning money, an exogenous attribution would be implied, whereas if the purpose had been also identified as helping his mother, the attribution would be endogenous. Conversely, conceptualizing the action as earning money would lead to an endogenous attribution, if the actor's end has been similarly identified, whereas if the end had been identified as helping his mother, the action would appear to be exogenously motivated.

But what may be the conditions for assigning a given identity to an action? The assumption that identity attributions are hy-

pothetical suggests the applicability of the discounting principle and of the attributional criteria discussed already in reference to the problem of causality. For instance, according to the discounting rule, the greater the number of hypotheses about an effect's (e.g., an action's) identity the lower will be the attributor's confidence in each such identity. Furthermore, a given identity hypothesis would be rendered more plausible the greater the amount of corroborative evidence in its favor. As in the case of causality, the evidence regarding identity could be assessed by reference to the criteria of consensus, consistency, or the convergence of information. In our earlier example, such criteria might more readily suggest the action's identity to be, say, sweeping the floor, than helping the actor's mother.

Earlier it has been implied that with an action's constant identity, endogeneity of attribution will vary with its perceived causality. The present discussion reverses this reasoning to suggest that with a constant causality (i.e., the action's purpose) the endogeneity of attribution will vary with this action's perceived identity. Thus, endogenous attribution will be rendered when the action is conceptualized in terms of its goal (i.e., as the attainment thereof) and exogenous attribution when it is conceptualized in terms that are foreign to its goal. Accordingly, endogenous attribution will be more likely when (a) no such foreign identity hypotheses will be entertainable and (b) corroborating evidence will imply the action's identity to be endogenous to the assumed goal.

The endogeneity question. The conditions of endogenous-exogenous attributions identified thus far involved modifying the attributor's conception of the action and its cause so as to vary the perceived similarity between the two. But how may one affect the likelihood of an endogenous attribution when the identities of both the action and the cause are relatively constant and refractory to change? In such a case a category of relevant conditions seems implied in the relation of endogeneity or exogeneity that may link action and goal. Specifically, the

endogeneity concept connotes the action's inherence to or inseparability from its goal. Thus, given a dissimilarly conceptualized action and goal, one could promote endogenous attribution by suggesting that the former is an integral part of the latter. For example, the young man who believes that only by dint of superior performance will he achieve the goal of rising in the company's ranks may become more endogenously motivated to his work than his counterpart who also contemplates alternative ways of attaining this objective (e.g., having good connections, marrying the boss's daughter).

RECAPITULATION AND CONCLUSIONS

The central thesis of this paper has been that the partition between endogenous and exogenous reasons is of considerable importance in the lay explanation of actions. A theoretical framework has been developed whereby the action's endogenous attribution implies the actor's intrinsic motivation and subjective freedom. In addition, with knowledge of the action's content, endogenous attribution has been shown to imply the action's specific intention. The present endogenous-exogenous distinction has been proposed to replace the frequent alternative partition of the actions' causes into the internal and external categories. It has been suggested that in lay assumption, actions are determined by the will, which is internal. In contrast, occurrences, which are (at least in part) independent of the will, are therefore explicable by attributes that may be internal or external.

The endogenous-exogenous and the internal-external frameworks have been compared on their degrees of conceptual coherence and empirical support in the domain of action explanation. It has been noted that the current applications of the internal-external distinction may suffer from two conceptual difficulties: (a) the problem of criterion and (b) the problem of inference. The first of these problems is that the classification of specific causes into the internal and external categories has typically lacked explicit justification, and the two criteria

that seem to underlie such categorizations (the notions of generically internal or external causes and of the uniqueness standard) need not coincide. The second problem is that the existing treatments of the internal-external distinction furnish little rationale for linking it with the various inferences with which it has been credited. Yet numerous examples may be imagined where the alleged inferences from internal versus external attributions (e.g., as interpreted in terms of the uniqueness standard) would seem a priori unlikely. By contrast, the endogenous-exogenous partition seems clear of the problems just outlined. The conceptual criterion for this distinction derives from the common differentiation between ends and means, and the inferences implied by the endogenous-exogenous partition are explicitly deducible therefrom, as shown in an earlier section of this paper.

On the empirical level, it has been noted that whenever research findings on actions have been explained by reference to the internal-external distinction, a reinterpretation was possible in terms of the endogenous-exogenous distinction. Furthermore, specific experimental comparisons yielded consistent support for the endogenous-exogenous distinction and failed to support the internal-external one.

The present article suggests several possible avenues for further research. To sample just a few:

1. Additional data seem necessary to demonstrate that internal attributions (uniqueness, self-selection, etc.) may not provide the explanatory implications with which they have been credited in the numerous areas of attributional research.

2. The broad claims connected with the endogenous-exogenous partition (of action's reasons) merit a careful empirical study. For instance, the contention that endogenous attribution is the necessary and sufficient condition of perceived freedom needs to be examined against the recently growing literature on the freedom notion. Similarly, the suggestion that inferences about irrational behaviors may be explicable by reference to the endogenous-exogenous distinction needs empirical substantiation.

3. Finally, the presently proposed conditions of endogenous versus exogenous attribution require a close investigation. It is by thorough understanding of such conditions that the theory of endogenous attribution may find its application in concrete situations of practical significance.

REFERENCE NOTE

1. Kruglanski, A. W. *Inferring enjoyment and freedom from an endogenous attribution of actions: A critique of the internal-external distinction*. Unpublished manuscript, University of North Carolina, Chapel Hill, 1974.

REFERENCES

- Ajzen, I., & Fishbein, M. A Bayesian analysis of attribution processes. *Psychological Bulletin*, 1975, 82, 261-277.
- Bandler, R. J., Madaras, G. R., & Bem, D. J. Self-observation as a source of pain perception. *Journal of Personality and Social Psychology*, 1968, 9, 205-209.
- Bem, D. J. An experimental analysis of self-persuasion. *Journal of Experimental Social Psychology*, 1965, 1, 199-218.
- Bem, D. J. Self-perception: An alternative interpretation of cognitive dissonance phenomena. *Psychological Review*, 1967, 74, 183-200.
- Bem, D. J. Self-perception theory. In L. Berkowitz (Ed.), *Advances in experimental social psychology* (Vol. 6). New York: Academic Press, 1972.
- Campbell, D. T. Social attitudes and other acquired behavioral dispositions. In S. Koch (Ed.), *Psychology: A study of a science* (Vol. 6). New York: McGraw-Hill, 1963.
- Cooper, J., Jones, E. E., & Tuller, S. M. Attribution, dissonance and the illusion of uniqueness. *Journal of Experimental Psychology*, 1972, 8, 45-57.
- deCharms, R. *Personal causation*. New York: Academic Press, 1968.
- Deci, E. L. Effects of externally mediated rewards on intrinsic motivation. *Journal of Personality and Social Psychology*, 1971, 18, 105-115.
- Deci, E. L. The effects of contingent and non-contingent rewards and controls on intrinsic motivation. *Organizational Behavior and Human Performance*, 1972, 8, 217-229. (a)
- Deci, E. L. Intrinsic motivation, extrinsic reinforcement, and inequity. *Journal of Personality and Social Psychology*, 1972, 22, 113-120. (b)
- Deci, E. L. *Intrinsic motivation*. New York: Plenum Press, 1974.
- Eacker, J. N. On some elementary philosophical problems of psychology. *American Psychologist*, 1972, 27, 553-566.
- Festinger, L., & Carlsmith, J. M. Cognitive consequences of forced compliance. *Journal of Abnormal and Social Psychology*, 1959, 58, 203-210.

- Goranson, R. E., & Berkowitz, L. Reciprocity and responsibility reactions to prior help. *Journal of Personality and Social Psychology*, 1966, 3, 227-232.
- Greene, D., & Lepper, M. R. How to turn play into work. *Psychology Today*, September 1974, pp. 49-54.
- Jones, E. E., & Davis, K. E. From acts to dispositions: The attribution process in person perception. In L. Berkowitz (Ed.), *Advances in experimental social psychology* (Vol. 2). New York: Academic Press, 1965.
- Jones, E. E., Davis, K. E., & Gergen, K. J. Role playing variations and their informational value for person perception. *Journal of Abnormal and Social Psychology*, 1961, 63, 302-310.
- Jones, E. E., Kanouse, D. E., Kelley, H. H., Nisbett, R. E., Valins, S., & Weiner, B. (Eds.), *Attribution: Perceiving the causes of behavior*. Morristown, N.J.: General Learning Press, 1972.
- Kelley, H. H. Attribution theory in social psychology. In D. Levine (Ed.), *Nebraska Symposium on Motivation*. Lincoln: University of Nebraska Press, 1967.
- Kelley, H. H. *Attribution in social interaction*. Morristown, N. J.: General Learning Press, 1971. (a)
- Kelley, H. H. *Causal schemata and the attribution process*. Morristown, N.J.: General Learning Press, 1971. (b)
- Kelley, H. H. The processes of causal attribution. *American Psychologist*, 1973, 28, 107-128.
- Kiesler, C. A., Nisbett, R. E., & Zanna, M. P. On inferring one's beliefs from one's behavior. *Journal of Personality and Social Psychology*, 1969, 11, 321-327.
- Kruglanski, A. W., Alon, S., & Lewis, T. Retrospective misattribution and task enjoyment. *Journal of Experimental Social Psychology*, 1972, 8, 493-501.
- Kruglanski, A. W., & Cohen, M. Attributed freedom and personal causation. *Journal of Personality and Social Psychology*, 1973, 26, 245-250.
- Kruglanski, A. W., Friedman, I., & Zeevi, G. The effects of extrinsic incentive on some qualitative aspects of task performance. *Journal of Personality*, 1971, 39, 606-617.
- Kruglanski, A. W., Riter, A., Amitai, A., Margolin, B. Shabtai, L., & Zaksh, D. Can money enhance intrinsic motivation? A test of the content-consequence hypothesis. *Journal of Personality and Social Psychology*, 1975, 31, 744-750.
- Lepper, M. R., Greene, D., & Nisbett, R. E. Undermining children's intrinsic interest in the extrinsic rewards: A test of the overjustification hypothesis. *Journal of Personality and Social Psychology*, 1973, 28, 129-137.
- Nisbett, R. E., & Schachter, S. Cognitive manipulation of pain. *Journal of Experimental Social Psychology*, 1966, 2, 227-236.
- Nisbett, R. E., & Valins, S. *Perceiving the causes of one's own behavior*. Morristown, N.J.: General Learning Press, 1971.
- Skinner, B. F. *Verbal behavior*. New York: Appleton-Century-Crofts, 1957.
- Steiner, I. D. Perceived freedom. In L. Berkowitz (Ed.), *Advances in experimental social psychology* (Vol. 5). New York: Academic Press, 1970.
- Thibaut, J. W., & Riecken, H. W. Some determinants and consequences of the perception of social causality. *Journal of Personality*, 1955, 24, 113-133.
- Weiner, B., Frieze, I., Kukla, A., Reed, L., Rest, S., & Rosenbaum, R. M. *Perceiving the causes of success and failure*. Morristown, N.J.: General Learning Press, 1971.

(Received April 11, 1975)

Announcement

The Publications and Communications Board of the American Psychological Association announces the appointment of William K. Estes as editor of *Psychological Review* for the years 1977 to 1982.

A Spreading-Activation Theory of Semantic Processing

Allan M. Collins
*Bolt Beranek and Newman Inc.,
Cambridge, Massachusetts*

Elizabeth F. Loftus
University of Washington

This paper presents a spreading-activation theory of human semantic processing, which can be applied to a wide range of recent experimental results. The theory is based on Quillian's theory of semantic memory search and semantic preparation, or priming. In conjunction with this, several of the misconceptions concerning Quillian's theory are discussed. A number of additional assumptions are proposed for his theory in order to apply it to recent experiments. The present paper shows how the extended theory can account for results of several production experiments by Loftus, Juola and Atkinson's multiple-category experiment, Conrad's sentence-verification experiments, and several categorization experiments on the effect of semantic relatedness and typicality by Holyoak and Glass, Rips, Shoben, and Smith, and Rosch. The paper also provides a critique of the Smith, Shoben, and Rips model for categorization judgments.

Some years ago, Quillian¹ (1962, 1967) proposed a spreading-activation theory of human semantic processing that he tried to implement in computer simulations of memory search (Quillian, 1966) and comprehension (Quillian, 1969). The theory viewed memory search as activation spreading from two or more concept nodes in a semantic network until an intersection was found. The effects of preparation (or priming) in semantic memory were also explained in terms of spreading activation from the node of the primed concept. Rather than a theory to explain data, it was a theory designed to show how to build human semantic structure and processing into a computer.

This research was supported by the National Institute of Education, U.S. Department of Health, Education, and Welfare, Project 1-0420, under Contract OEC-1-71-0100(508) with Bolt Beranek and Newman Inc. Revision of the paper was supported by a grant from the John Simon Guggenheim Memorial Foundation to the first author. We would like to thank Stephen Woods, Colin MacLeod, Mark L. Miller, and the reviewers for their comments on previous drafts of this paper.

Requests for reprints should be sent to Allan M. Collins at Bolt Beranek and Newman Inc., 50 Moulton Street, Cambridge, Massachusetts 02138.

¹ Quillian's theory of priming appeared in the unpublished version of the 1967 paper (i.e., CIP Paper No. 79, Carnegie Institute of Technology, 1965).

Since the theory was proposed, there have been a number of experiments investigating retrieval and priming in semantic memory. In the present paper, we attempt to show how an elaboration of Quillian's basic theory can account for many of the results. In the first section, we briefly review the original theory while trying to correct a number of the common misunderstandings concerning it. In the second section, we extend the theory in several respects, and in the third section show how the extended theory deals with some recent experimental findings. In the fourth section we compare the theory to the model of Smith, Shoben, and Rips (1974).

QUILLIAN'S THEORY OF SEMANTIC MEMORY

The fact that Quillian's theory was developed as a program for a digital computer imposed certain constraints on the theory, which Quillian felt were psychologically unrealistic. We will recount the theory as he proposed it, and then elaborate the theory in psychological terms. The theory made a number of assumptions about structure and processing in human semantic memory. A brief discussion of these assumptions follows.

People's concepts contain indefinitely large amounts of information. Quillian used the example of a machine. If one asks people to tell everything they know about machines,

they will start off giving obvious properties, for example, that machines are man-made and have moving parts. But soon people run out of obvious facts and begin giving facts that are less and less relevant, for example, that a typewriter is a machine or even that the keys on the IBM electric typewriters select the position of a ball that strikes the ribbon against the paper. The amount of information a person can generate about any concept in this way seems unlimited.

In these terms concepts correspond to particular senses of words or phrases. For example, not only is the noun "machine" a concept, but the verb "to machine" is a concept; the "particular old car I own" is a concept; the notion of "driving a car" is a concept; even the notion of "what to do if you see a red light" has to be a concept. Thus, people must have a very large number of concepts, and concepts must have very complicated structures.

A concept can be represented as a node in a network, with properties of the concept represented as labeled relational links from the node to other concept nodes. These links are pointers, and usually go in both directions between two concepts. Links can have different *criticalities*, which are numbers indicating how essential each link is to the meaning of the concept. The criticalities on any pair of links between two concepts can be different; for example, it might be highly critical for the concept of a typewriter that it is a machine, and not very critical for the concept of machine that one kind is a typewriter. From each of the nodes linked to a given node, there will be links to other concept nodes and from each of these in turn to still others. In Quillian's theory, the full meaning of any concept is the whole network as entered from the concept node.

The links are not simply undifferentiated links, but must be complicated enough to represent any relation between two concepts. In the original theory, Quillian proposed five different kinds of links: (a) superordinate ("isa") and subordinate links, (b) modifier links, (c) disjunctive sets of links, (d) con-

junctive sets of links, and (e) a residual class of links, which allowed the specification of any relationship where the relationship (usually a verb relationship) itself was a concept. These different kinds of links could be nested or embedded to any degree of depth, so that the format was designed to be flexible enough to express anything, however vague or specific, that can be expressed in natural language.

The search in memory between concepts involves tracing out in parallel (simulated in the computer by a breadth-first search) along the links from the node of each concept specified by the input words. The words might be part of a sentence or stimuli in an experimental task. The spread of activation constantly expands, first to all the nodes linked to the first node, then to all the nodes linked to each of these nodes, and so on. At each node reached in this process, an activation tag is left that specifies the starting node and the immediate predecessor. When a tag from another starting node is encountered, an *intersection* between the two nodes has been found.² By following the tags back to both starting nodes, the *path* that led to the intersection can be reconstructed.

When an intersection has been found, it is necessary to *evaluate* the path to decide if it satisfies the constraints imposed by syntax and context. The complicated kinds of decision rules that are invoked for comprehension of sentences in this evaluation phase are described in Quillian (1969). For categorization tasks these rules are described by Collins and Quillian (1972b) and in the next section of this paper. As an example, in a phrase such as "the fall leaves," a path found between the concept "to fall" and the

² There can be intersections with more than two starting nodes, but we have limited our discussion in this paper to the case of two nodes (as did Quillian, 1966, initially). The basic assumptions in Quillian's theory and our elaboration can apply to intersections with more than two starting nodes, but this leads to complications in the evaluation of intersections. Some of these were discussed by Quillian (1969) in regard to comprehension, but for the experiments considered in this paper, only the case of two starting nodes needs to be considered.

concept "tree leaf" would be rejected as a wrong interpretation, because syntax requires a participial form of "fall" to fit that interpretation. If a path found is rejected, other paths are considered in the order in which they are found.

Priming (or preparation) involves the same tracing process that was described for memory search. When a concept is primed, activation tags are spread by tracing an expanding set of links in the network out to some unspecified depth. When another concept is subsequently presented, it has to make contact with one of the tags left earlier to find an intersection. One of the non-obvious implications of this view of priming is that links as well as nodes will be primed. This is because Quillian treated links themselves as concepts (see above). Thus priming a node such as "red" will prime the links involving the relation "color" throughout the network. This provides a very powerful context mechanism.

Common Misinterpretations of Quillian's Theory

There is a rich variety of misinterpretations of Quillian's theory, many of them deriving from Collins's (Collins & Quillian, 1969, 1970a, 1970b) simplifications of the theory. The problem arose because Collins and Quillian were investigating specific aspects of the theory and only described enough of the theory to motivate a particular experiment. In turn experimenters made interpretations of these simplified versions, which did not fit with the theory as described elsewhere (Bell & Quillian, 1971; Quillian, 1966, 1969).

Perhaps the most prevalent misinterpretation of Quillian's theory concerns the idea of cognitive economy (Anderson & Bower, 1973; Conrad, 1972). In this regard, it is important to distinguish the strong theory of cognitive economy, which Conrad takes issue with in her attack on Collins and Quillian (1969), and the weak theory of cognitive economy, which Collins and Quillian were testing (though they did not spell it out clearly enough). As Conrad (1972) states, she rejects the "hypothesis that all proper-

ties are stored only once in memory and must be retrieved through a series of inferences for all words except those that they most directly define" (p. 153). This is a statement of the strong theory of cognitive economy. Undoubtedly the Collins and Quillian (1969) paper gave rise to this notion, but the authors cautioned against making that interpretation of the theory. As they said, "people surely store certain properties at more than one level in the hierarchy" (p. 242), and they cited the maple leaf as an example of this general rule.

The strong theory requires erasing information whenever it applies at a more general level. If a person learns a robin can fly and then later that birds fly, the strong theory implies that "flying" must be erased from "robin." The weak theory of cognitive economy merely assumes that every time one learns that X is a bird, one does not at that time store all the properties of birds with X in memory. Thus, an inference will be necessary to decide that X can fly, unless one encounters this fact directly. Hence, Collins and Quillian, in testing the weak theory of cognitive economy, picked instances where people were not likely to have encountered the general property with the specific instance (e.g., "A wren can fly"). The point of the experiment was to test whether it was possible to measure inference time, when the weak theory of cognitive economy implies that an inference is likely to be necessary for most subjects.

Another assumption sometimes made about Quillian's theory is that all links are equal (Anderson & McGaw, 1973; Rips, Shoben, & Smith, 1973; Wilkins, Note 1). In Quillian's original theory, there were criteriality tags on links, as we described earlier. In Collins and Quillian (1969, 1972b) links were assumed to have differential accessibility (i.e., strength or travel time). The accessibility of a property depends on how often a person thinks about or uses a property of a concept. Whether criteriality and accessibility are treated as the same or different is a complex issue, but network models allow them to be treated either way. Thus for example, even though "lungs,"

"hands," and "warts" are all linked directly to the concept "human," these links need not be in any sense equal. The same is true for the links between "bird" and its exemplars, such as "robin," "chicken," or "penguin." Rips et al. (1973) suggest that intermediate nodes are necessary for a network model to explain the reaction time differences they find in categorizing different birds. This makes the mistaken assumption that all links are equally criterial or accessible in any network model. It turns out, however, that differences in links are crucial to many different aspects of human semantic processing as Carbonell and Collins (1973) point out in their discussion of importance (or criteriality) tags.

A related implication of the Rips et al. (1973) paper and also a more recent paper of Smith et al. (1974) is that feature models can account for data that network models cannot. A feature model posits that a concept consists of a set of values on a large number of semantic dimensions (e.g., animateness, color, etc.). What is strange about this argument is that network models were developed as a method of representing features in a computer. Any process that can be represented in a feature model is representable in a network model; in particular, the Smith et al. model itself could be implemented in a semantic network (Hollan, 1975). In fact, network models are probably *more* powerful than feature models, because it is not obvious how to handle inferential processing or embedding in feature models.

Smith et al. (1974) argued in favor of feature models because their data for comparison of concepts seemed to fit a feature comparison process. What should be emphasized about Quillian's theory is that the parallel search would inevitably lead to just such a feature comparison process, though the process would take place over a period of time as different connections are found. One way that Quillian's theory is different from the Smith et al. models is that superordinate connections, if they exist, would also be found and evaluated. The distinction between these two theories is so crucial

that we will discuss it at length in conjunction with the spreading activation theory's explanation of the Rips et al. (1973) results.

Another common misconception of Quillian's theory shows up in Juola and Atkinson's (1971) work on categorization judgments. In a categorization task, response time is measured for a subject to decide whether or not a particular instance (e.g., "car") is a member of one or more categories (e.g., "flower" or "vehicle"). Juola and Atkinson assume that in Quillian's theory the memory search to make a categorization judgment proceeds from the instance to the category. In fact, the wording in Collins and Quillian (1969) mistakenly gives that impression. But Quillian's theory (1966, 1969) assumes the search proceeds from both the instance and category in parallel. However, if one or the other is presented first, this gives the search from that node a head start, which is the notion of priming. Juola and Atkinson's experiment involves priming in a complicated way, which we will discuss below.

Anderson and Bower (1973) reject a Quillian-like model of a parallel search, while acknowledging that their data are compatible with "a parallel model whose search rate is slower in proportion to the number of paths that must be searched" (p. 371). Anderson and Bower's argument implies wrongly that Quillian has made the independence assumption for his parallel search. An independent parallel search is like a race where the speed of each runner is independent of the other runners. This is a common assumption in psychology, because it makes it possible to assign an upper bound to reaction time (see Sternberg, 1966). But there is no difficulty for Quillian's theory if the parallel search rate depends on the number of paths searched. Hence, Anderson and Bower's data are perfectly compatible with Quillian's parallel search.

The above discussion, then, shows what Quillian's theory is *not*, or at least some of what it is not. Several other misconceptions are discussed in Collins and Quillian (1972b), in particular the notion that Quil-

lian's theory of memory is rigidly hierarchical, which Anderson and Bower (1973, p. 379) still believe, and Schaeffer and Wallace's (1970) argument that Quillian's theory predicts it will always take less time to compare concepts that are close together in the semantic network. We will return to some of these same papers below, in order to describe how the extended version of Quillian's theory accounts for some of the results these experimenters have used to reject Quillian's theory.

THE EXTENDED THEORY

In order to deal with the specific experimental results that have appeared in recent years, several more processing and structural assumptions must be added to the basic Quillian theory. These do not bend the theory, but merely elaborate it in such a way that it can be applied to the kinds of experiments on semantic memory that have been performed recently. The elaboration may itself be wrong, so our mistakes should not be held against Quillian's theory.

Local Processing Assumptions

There are four local processing assumptions in the extended theory. These four assumptions transform the theory from computer terms to quasi-neurological terms, à la Pavlov. But all the assumptions of the original theory should be preserved despite the transformation, except that activation tags are to be considered as source-specific activation (i.e., activation that is traceable to its node of origin).

1. When a concept is processed (or stimulated), activation spreads out along the paths of the network in a decreasing gradient. The decrease is inversely proportional to the accessibility or strength of the links in the path. Thus, activation is like a signal from a source that is attenuated as it travels outward.

2. The longer a concept is continuously processed (either by reading, hearing, or rehearsing it), the longer activation is released from the node of the concept at a fixed rate. Only one concept can be actively processed at a time, which is a limita-

tion imposed by the serial nature of the human central process (Collins & Quillian, 1972b). This means that activation can only start out at one node at a time. But it continues in parallel from other nodes that are encountered as it spreads out from the node of origin.

3. Activation decreases over time and/or intervening activity. This is a noncommittal assumption that activation goes away gradually by some mechanism. Assumptions 2 and 3 impose a limitation on the amount of activation that can be allocated in priming more than one concept, because the more concepts that are primed, the less each will be primed.

4. With the assumption that activation is a variable quantity, the notion of intersection requires a threshold for firing. The assumption is that activation from different sources summates and that when the summation at the point of intersection reaches threshold, the path in the network producing the intersection will be evaluated.

Global Assumptions About Memory Structure and Processing

There are three assumptions in the extended theory concerned with the global structure of memory and its processing. These are generalizations of Loftus's (Note 2) arguments that semantic memory is organized primarily into noun categories and that there is a "dictionary" (or lexical memory) separate from the conceptual network.

5. The conceptual (semantic) network is organized along the lines of semantic similarity. The more properties two concepts have in common, the more links there are between the two nodes via these properties and the more closely related are the concepts. This means that different vehicles or different colors will all be highly interlinked through their common properties. This also implies that red things (e.g., fire engines, cherries, sunsets, and roses) are *not* closely interlinked, despite the one property they have in common. In these terms semantic relatedness is based on an aggrega-

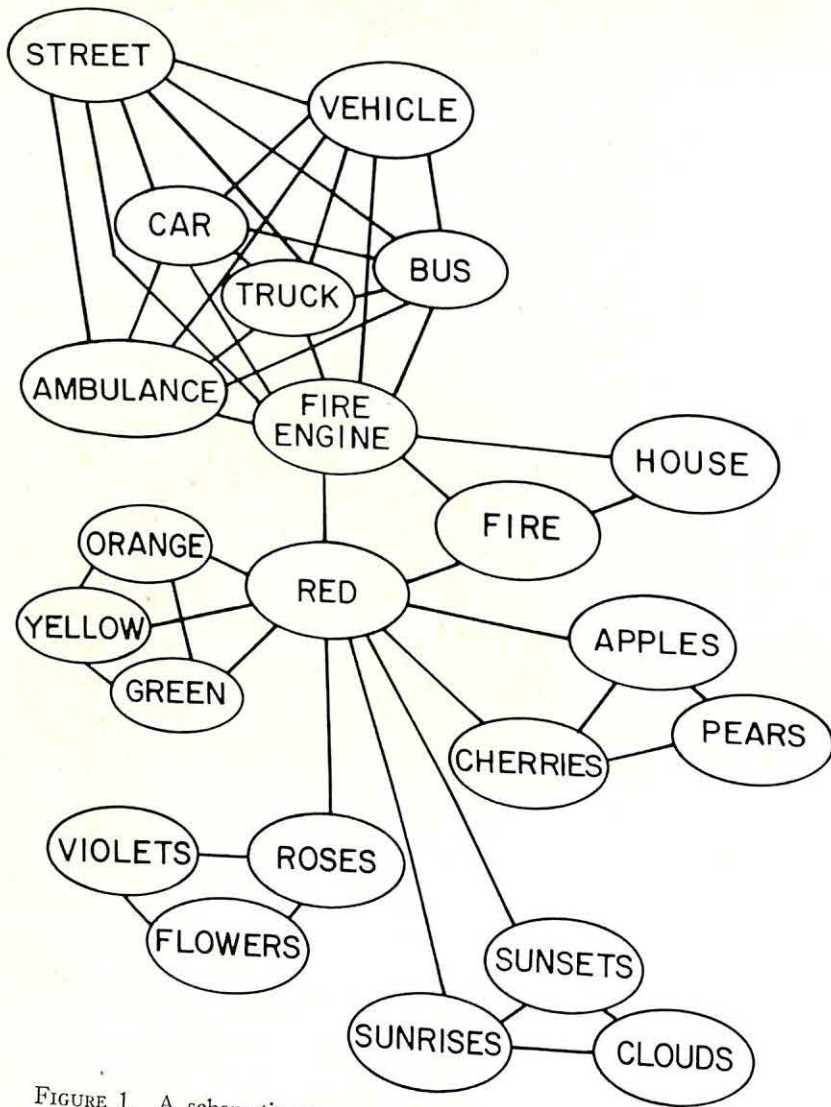


FIGURE 1. A schematic representation of concept relatedness in a stereotypical fragment of human memory (where a shorter line represents greater relatedness).

gate of the interconnections between two concepts.³

³ Semantic relatedness is a slightly different notion from semantic distance, though the two terms are sometimes used interchangeably. Semantic distance is the distance along the shortest path, and semantic relatedness (or similarity) is an aggregate of all the paths. Two concepts may be close in distance, say by a path through "red," and still not be closely related because that is the only path. Our use of *close* to refer to both relationships is admittedly confusing. In this paper we shall use *close* to refer to relatedness or similarity, though in some tasks (Quillian, 1966) it is only distance that matters.

Figure 1 illustrates this aggregate notion of concept relatedness for a hypothetical human memory. (It is the kind of diagram that the scaling techniques of Rips et al., 1973, would produce.) In the figure the various vehicles are shown as closely related, because of the numerous individual connections that are assumed to exist between them. Conversely, the concepts associated with "red" are shown as less related, because of the presumed paucity of interconnections between them.

From the assumption that memory is organized according to semantic similarity, to-

gether with earlier assumptions, it follows that if "vehicle" is primed, activation at any type of vehicle will accumulate from many neighboring nodes. That is to say, to the degree that "fire engine" is primed by "vehicle," it will in turn prime "ambulance," "truck," "bus," etc., and each of these in turn will prime the others. On the other hand, if "red" is primed, the activation that spreads to "fire engine" will not prime "cherries," "roses," or "sunsets" to any great extent, because there are so few connections between these concepts. Instead, "fire engine" will tend to prime other vehicles, and "cherries" to prime other fruits. Hence, the same amount of activation will be diffused among a greater number of concepts.

6. The names of concepts are stored in a lexical network (or dictionary) that is organized along lines of phonemic (and to some degree orthographic) similarity. The links from each node in the lexical network are the phonemic properties of the name, specified with respect to their position in the word. The properties stored about names are assumed to be the properties that Brown and McNeill (1966) found people could identify about words on the "tip of their tongue." Each name node in the lexical network is connected to one or more concept nodes in the semantic network.

7. Loftus's (Note 2) data lead to the further assumption that a person can control whether he primes the lexical network, the semantic network, or both. For example, a person can control whether to prime (a) words in the lexical network that sound like "bird," (b) concepts in the semantic network related to "bird," or (c) words in the lexical network corresponding to the concepts in (b). This control over priming can be thought of in terms of summation of diffuse activation for an entire network (perhaps in a particular part of the brain) and source-specific activation released from a particular node. Thus, (a) would derive from activation of the lexical network together with the word "bird," (b) would derive from activation of the semantic network together with the concept "bird," and (c)

would derive from activation of both networks together with the concept "bird."

Assumptions About Semantic Matching Process

There are a number of assumptions about the decision process for evaluating whether or not two concepts match semantically. This is a fundamental process that occurs in many aspects of language processing, such as matching referents, assigning cases, and answering questions (Collins & Quillian, 1972b; Collins, Warnock, Aiello, & Miller, 1975). Categorization tasks, which ask "Is X a Y?" (where X and Y are concepts), directly investigate this process. The decision process described here is a more explicit and somewhat revised version of the process postulated by Collins and Quillian (1972b), with additions to encompass the results of Holyoak and Glass (1975).

8. In order to decide whether or not a concept matches another concept, enough evidence must be collected to exceed either a positive or a negative criterion. The evidence consists of various kinds of intersections that are found during the memory search. Evidence from different paths in memory sum together. Positive and negative evidence act to cancel each other out, as shown by dialogue excerpts in Carbonell and Collins (1973). Failure to reach either criterion before running out of relevant evidence leads to a "don't know" response (Collins et al., 1975). This process is essentially the Bayesian decision model that is common in the reaction time literature (see, for example, Fitts, 1966; Stone, 1960).

There are a number of different kinds of paths between the two concepts that constitute positive or negative evidence. Any of these types of evidence might contribute to a particular decision. The different types are listed in Table 1 and described below in Assumptions 9-13.

9. If the memory search finds that there is a superordinate (or a negative superordinate) connection from X to Y, that fact alone can push the decision over the positive (or negative) criterion. Superordinate links act like highly criterial property links

TABLE 1
TYPES OF PATHS FOUND IN MEMORY
THAT CONSTITUTE POSITIVE OR
NEGATIVE EVIDENCE

Positive evidence	Negative evidence
Superordinate connection	Negative superordinate connection
Property comparison, matching property	Property comparison, distinguishing property
Wittgenstein strategy, matching property	Wittgenstein strategy, distinguishing property
	Mutually exclusive subordinates
	Counterexamples

(see below). For example, it is conclusive positive evidence that a mallard is a bird, if superordinate links are found between "mallard" and "duck" and between "duck" and "bird." Similarly, if a negative superordinate link is found between "bat" and "bird," it is conclusive evidence that a bat is not a bird.

10. If the memory search finds properties on which X and Y match (i.e., common properties), this is positive evidence proportional to the criteriality of the property for Y. If the memory search finds properties on which X and Y mismatch (i.e., distinguishing properties), this is negative evidence proportional to the criteriality of the property for Y. There is an asymmetry in the weighing of positive and negative evidence in a property comparison, because a mismatch on just one fairly criterial property can lead to a negative decision, whereas most of the highly criterial properties must match in order to reach a positive decision (Collins & Quillian, 1972b).

It is important to note that property comparisons and superordinate connections sum together in reaching either criterion as the memory search finds them. Thus, distinguishing properties make it harder to reach the positive criterion when there is a superordinate connection and therefore slow down the process.

As an example of property comparison, suppose there is no superordinate connection in a particular person's memory between "mink" and "farm animal" and between "cat" and "farm animal." Then the decision as to whether minks or cats are farm

animals might be based on a comparison of the properties of minks or cats on one hand, and farm animals on the other. The most criterial properties of farm animals are presumably being animate and being kept on farms, but other less criterial properties include being domesticated, being raised for some purpose, or being kept in barns or outside. How a particular person would weigh the various properties of minks and cats to decide whether they are farm animals would vary from person to person (our intuition is that minks are farm animals and cats are not, even though both have the two most criterial properties of farm animals—what Smith et al., 1974, called defining properties). This decision strategy is similar to that proposed by Smith et al., as we will discuss later.

11. The Wittgenstein strategy is a variant of the property comparison strategy. It is postulated on the basis of Wittgenstein's (1953) observation that to decide whether something is a game (for example, frisbee), a person compares it to similar instances that are known to be games. Our assumption is that if any properties of X are found that match properties of another instance whose superordinate is Y, these constitute positive evidence. Similarly, any distinguishing properties constitute negative evidence. In the Wittgenstein strategy, unlike the property comparison strategy, matching properties count just as much toward a positive decision as distinguishing properties count toward a negative decision.

To illustrate the Wittgenstein strategy, Collins and Quillian (1972b) pointed out that in deciding whether a stagecoach is a vehicle, it might be compared to a car. The many properties that a stagecoach has in common with a car constitute strong positive evidence that a stagecoach is a vehicle. But notice that a stagecoach does not have a motor, which is highly criterial for being a car. Though this is strong evidence that a stagecoach is not a car, it is only weak evidence that it is not a vehicle. This illustrates how the same evidence is weighed differently in the property comparison strategy and the Wittgenstein strategy. The

final decision that a stagecoach is a vehicle might depend both on matching properties between a stagecoach and vehicles in general (conveyance, motion, etc.) and matching properties between a stagecoach and particular vehicles like a car (seats, doors, etc.). Thus the property comparison strategy and the Wittgenstein strategy might combine to determine a person's response.

12. The mutually exclusive subordinates strategy was necessary for programming a computer to answer questions (Collins et al., 1975). Holyoak and Glass (1975) argue that this strategy accounts for some of their reaction time data (they call it a contradiction). The assumption is that if two concepts have a common superordinate with mutually exclusive links into the common superordinate, then this constitutes strong negative evidence, almost comparable to a negative superordinate link.

For example, if the question is whether a mallard is an eagle, the fact that a mallard is a duck and ducks and eagles are mutually exclusive kinds of birds is rather conclusive evidence that a mallard is not an eagle. Though Holyoak and Glass (1975) do not mention it, the mutually exclusive restriction is necessary. For example, the fact that Mike Mansfield is a politician does not exclude him from being a lawyer. Although "politician" and "lawyer" are both occupational roles, they are not mutually exclusive and in fact most politicians are lawyers. But lacking specific information to the contrary, people may make a default assumption of mutual exclusivity when two concepts have a common superordinate.

13. Counterexamples also can be used as negative evidence. This strategy derives from Holyoak and Glass (1975), who argue that statements of the kind "All birds are canaries" are disconfirmed by finding a counterexample, such as "robin." If the question is of the form "Is X a Y?" and there is a superordinate link from Y to X, then finding a counterexample involves finding a Z that also has X as superordinate and is mutually exclusive from Y. This is conclusive evidence that X is not always a Y.

Holyoak and Glass (1975) discuss counterexamples in the context of the universal quantifier "all," but the same process would occur for a question of the kind "Is a marsupial a kangaroo?" In such a case, retrieving a counterexample (such as a wallaby) can be used to determine that kangaroos are a subset of marsupials and not equivalent (e.g., "automobiles" and "cars" are equivalent concepts).

Though these five kinds of evidence (Assumptions 9-13) are the only ones we have postulated for the semantic matching process seen in categorization tasks, there may be other kinds of evidence of this sort. We should stress that there are many other kinds of evidence people use for answering more complicated questions (Collins et al., 1975). It is beyond the scope of this paper, however, to consider all the different ways people use evidence to make semantic decisions.

RECENT EXPERIMENTS

In this section, we discuss how the theory deals with some different kinds of recent experiments. The four types of studies to which we apply the theory are (a) several production experiments by Loftus (Freedman & Loftus, 1971; Loftus, 1973a, 1973b, Note 2); (b) Juola and Atkinson's (1971) multiple-category experiment; (c) the Conrad (1972) sentence-verification experiment; and (d) several categorization experiments on the effects of semantic relatedness and typicality (Holyoak & Glass, 1975; Rips et al., 1973; Rosch, 1973; Smith et al., 1974). We intend to deal with the major kinds of available findings to which the Quillian theory has not yet been applied. Our objective is to show how a spreading-activation theory can handle these results, not to consider all the possible alternative explanations of the experiments.

Production Experiments of Loftus

There are several Loftus experiments we want to discuss in terms of the spreading-activation theory. The first of these is an experiment by Freedman and Loftus (1971), in which subjects had to produce an in-

stance of a category that began with a given letter or was characterized by a given adjective. For example, subjects might be asked to name a fruit that begins with the letter *A* or a fruit that is red. On some trials the category was shown first and on some trials second. Hence, this was a priming experiment in that one concept was activated before the other. Reaction time was measured from the onset of the second stimulus.

Our concern is with the finding that subjects were faster when the category (e.g., "fruit") was given first than when either the letter or the adjective was given first. This basic result was later replicated even for cases in which the instance named was a more frequent associate to the adjective than to the category noun (e.g., "lemon" is a closer associate of "sour" than of "fruit").

The explanation in terms of the theory is as follows: When a noun, such as "fruit," is presented first, the activation spreads to nodes connected to "fruit," among which are instances such as "apple," "pear," "peach," "orange," and "lemon." But these concepts are all highly interlinked with each other (though some, such as "orange" and "lemon," are more closely interlinked than others). Thus, the total amount of activation is spread among a relatively small number of closely interlinked concepts (see Assumption 5). However, when an adjective or letter is presented first, say "red" or "A," the activation spreads to a much wider set of concepts, which are not particularly interlinked with each other. Thus, the large variety of different things that are red or that start with the letter *A* will receive relatively little priming when the adjective or letter are given first. Because priming the noun leads to a greater accumulation of activation on the instances, these are closer to their threshold for firing, so that it takes less stimulation, and hence less time, to trigger an intersection when the second stimulus is presented.

Freedman and Loftus (1971) explained their finding in terms of entering the category when a noun is presented and entering a cluster within the category when the adject-

tive or letter is presented. Thus if the noun is presented first, the subject can enter the category immediately and need only choose the correct cluster when the adjective or letter is presented. But if the adjective or letter is presented first, the subject must wait until the category is presented, because the cluster is specific to that category. (However, Loftus, Note 2, has revised this explanation for the letter stimulus in her dictionary-network model.)

The Freedman and Loftus explanation is not altogether different from the explanation offered here, though our theory is less rigidly hierarchical. The rigid hierarchy gets into trouble with errors such as one we encountered where a subject produced "Ben Franklin," given the stimulus pair "president" and "F," although he later recognized his mistake. In an activation theory, "Franklin" is a very likely intersection starting at "president" and "F," because he is so closely linked with the concept, "president," and some of its foremost instances, such as "Washington." Such a wrong intersection was likely in this case because the correct answer (prior to Ford) was "Fillmore," who is rather inaccessible and unlikely to be found quickly enough to preclude finding "Franklin." Once such an intersection is found, it is only by evaluating the connection between "president" and "Franklin," that it can be rejected (see Assumptions 8-13). It is a general problem of category-search models that they cannot deal with such errors.

Perhaps the major advantage of the spreading-activation theory over the Freedman and Loftus (1971) explanation is in tying their result to a parallel result in a quite different experiment by Loftus (1973b). In a categorization experiment, Loftus found that the direction of the association between the category and the instance determined whether subjects were faster when given the category first or the instance first. In the experiment she used four kinds of category-instance pairs: (a) pairs where both the category and instance evoked the other with high frequency (e.g., "tree-oak"); (b) pairs where the category evoked

the instance with high frequency, but the instance evoked the category with low frequency (e.g., "seafood-shrimp"); (c) pairs where the category evoked the instance with low frequency, but the instance evoked the category with high frequency (e.g., "insect-butterfly"); and (d) pairs where both the category and instance evoked the other with low frequency (e.g., "cloth-orlon"). When the category was presented before the instance, reaction time for Conditions (a) and (b) was approximately equal and significantly faster than for Conditions (c) and (d). However, when the instance was presented first, reaction time for Conditions (a) and (c) was approximately equal and significantly faster than for Conditions (b) and (d). That is to say, subjects are fast when the category is presented first, if the category evokes the instance with high frequency, and subjects are fast when the instance is presented first, if the instance evokes the category with high frequency. The spreading-activation theory explains the pattern of reaction times in the following way, assuming that production frequency is a measure of the strength or accessibility of the path from one concept to another. When the first concept (i.e., the one presented first) evokes the second with a relatively high frequency, this means that more activation spreads to the second, and it takes less time to reach the threshold for an intersection. Thus, the amount the first concept primes the second concept determines the reaction time.

By comparing this experiment with the Freedman and Loftus (1971) study, it can be seen that the two results are exactly parallel. Based on our structural assumptions, "fruit" primes "apple" more than "red" or the letter "A" primes "apple" in the Freedman and Loftus study. Hence, the shorter reaction time occurs when "fruit" or "A" is presented first. Similarly in the Loftus (1973b) study, when the category primes the instance most highly, the shortest reaction times occur when the category is presented first. But when the instance primes the category most highly, the shortest reaction times occur when the instance

is presented first. A spreading-activation explanation is quite compelling to account for the Loftus (1973b) results, and the theory offered here encompasses the order effect in both the Loftus study and the earlier Freedman and Loftus experiment within a single framework.

Recently Loftus (Note 2) has found two different ways in which presenting a letter acts differently from presenting an adjective, in variations of the Freedman and Loftus paradigm. This has led her to the development of a dictionary-network model, which we will translate into spreading-activation terms. The first difference between presenting a letter and an adjective appeared when Grober and Loftus (1974) compared reaction time in two conditions: one where noun-adjective (e.g., "fruit-red") and noun-letter (e.g., "fruit-A") trials were randomly intermixed, and one where noun-adjective and noun-letter trials were separated into blocks. In all cases the noun preceded the adjective or letter. The results of this experiment are shown in Figure 2. It is clear that when the subject knows a letter is coming, he can prepare for it. But in the mixed condition, the subject apparently prepares for either kind of trial the same way he prepares for an adjective trial, since adjective trials take the same amount of time in either case. The theory's description of semantic processing on the adjective trials is the same as that given earlier for the Freedman and Loftus (1971) experiment, with the amendment that only the semantic network and not the lexical network would be diffusely primed before the adjective is presented. When an intersection is found in the semantic network, then the subject must retrieve the name from the lexical network.

Loftus (Note 2) described what must happen on noun-letter trials in the blocked condition as follows:

The first step of the process is entering the category. The next step is a quasi-parallel simultaneous search towards the Dictionary. That is to say, the subject traces some number of pathways leading from category instances to the Dictionary representations of those instances. This step can be started during the interval between the

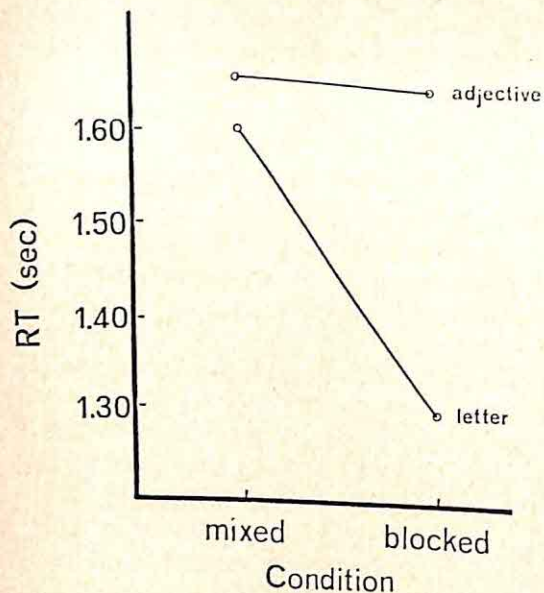


FIGURE 2. Reaction time for noun-adjective and noun-letter stimuli in mixed and blocked conditions (from Loftus, Note 2).

presentation of the category name and the restricting letter if the subject knows a letter is coming. (p. 13)

This is essentially the spreading-activation explanation, if the dictionary is taken to be a lexical network. Rather than saying that "the subject traces some number of pathways," which suggests a conscious tracing process, the present theory would say that activation spreads along some number of pathways, because the subject has activated the lexical network in addition to the semantic network (see Assumption 7). Hence, in the present explanation, the subject's control is reduced to diffusely activating whole networks rather than specific pathways (in addition to the specific nodes activated by the stimuli in the experiment). The difference in the results for the noun-letter trials in the two conditions then depends on whether the subject primes both networks (as in the blocked condition) or only the semantic network (as in the mixed condition). The reason he only activates the semantic network in the mixed condition may be either because of a principle of least effort (hence he could speed up his reaction time if he tried) or because there is less

activation available to the semantic network if both are primed (hence he will be slower on noun-adjective trials if he primes both).

As can be seen in Figure 2, the subject is much slower on noun-adjective trials than on noun-letter trials in the blocked condition. This is accounted for by the fact that an intersection on a noun-adjective trial occurs in the semantic network and requires the further step of retrieving the corresponding name in the lexical network. On the other hand, the intersection on a noun-letter trial occurs at the name in the lexical network. Therefore, the name does not then need to be retrieved.

The second result that shows up the difference between adjectives and letters was predicted by Loftus from the dictionary-network model. In this experiment (Loftus & Cole, 1974) subjects saw three stimuli, ordered either noun, adjective, letter or noun, letter, adjective. For example, the three stimuli might be "animal," "small," and "M," for which an appropriate response is "mouse." The prediction was that the subject should be faster when the adjective is presented before the letter, and this was the result found. The reasoning is as follows: When the adjective appears before the letter, activation will spread from a small set of instances in the semantic network to the lexical network where the intersection occurs, since the letter can be expected just as in the blocked condition. When the letter is presented before the adjective, activation will spread from a small set of instances in the lexical network back to the semantic network where an intersection with the adjective will occur. Then the subject must return again to the lexical network to retrieve the name, so there is an extra transit necessary in this condition.

Loftus has also run a series of experiments in which subjects were asked to produce a member of a category and a short time later asked to produce a different member of that category (Loftus, 1973a; Loftus, Senders, & Turkeltaub, 1974; Loftus & Loftus, 1974). This was accomplished by show-

ing a category-letter pair (e.g., "fruit-P"), which asked the subject for an appropriate instance, then, following 0, 1, or 2 intervening items, showing the same category paired with a different letter (e.g., "fruit-A"), which asked for a different instance. The general finding is that reaction time for the second instance is shorter than reaction time for the first instance and increases monotonically with the number of intervening items. For example, in Loftus (1973a) a subject's baseline time to name a fruit beginning with the letter "P" was 1.52 sec. However, it took him 1.22 sec to produce the same response if he had named a different fruit on the previous trial and 1.29 sec to produce the response if he had named a different fruit two trials back.

The spreading-activation theory predicts these results by assuming that when an item is processed, other items are activated to the extent that they are closely related to that item. That is, retrieving one category member produces a spread of activation to other category members, facilitating their later retrieval. The assumption (Assumption 3) that activation decreases over time or trials predicts the lag effect.

Meyer and Schvaneveldt (Meyer, 1973; Meyer & Schvaneveldt, 1971; Schvaneveldt & Meyer, 1973; Meyer, Schvaneveldt, & Ruddy, Note 3) have also shown that the time to retrieve information from memory is faster if related information has been accessed a short time previously. Their paradigm is somewhat different. Subjects were required to classify letter strings as words or nonwords. The general finding was that the response time to classify a letter string as a word is faster if the subject has just classified a semantically similar word as opposed to a semantically dissimilar word. Thus, for example, the time it takes to classify "butter" as a word is faster if "butter" is preceded by "bread" than if it is preceded by "nurse." Their results have led Meyer and Schvaneveldt to an explanation in terms of spreading activation and illustrate the widely different paradigms that such a theory can encompass.

Juola and Atkinson's Study with Multiple Categories

An increase in reaction time with multiple categories has been found by Juola and Atkinson (1971) in a task where subjects had to decide whether a stimulus word belonged to one of a variable number (1-4) of pre-specified (target) categories. They compared this task with one where subjects decided if the stimulus word was the same as one of a variable number (1-4) of target words. Their experiment was designed to distinguish between two kinds of models, one they attribute to Landauer and Freedman (1968) and one they attribute to Collins and Quillian (1970a). In most respects, their results fit the model they derived from Landauer and Freedman, but since the spreading-activation theory provides an alternative explanation for their results, we want to compare their two models with our theory.

The model Juola and Atkinson (1971) derived from Landauer and Freedman (1968) is very similar to what Landauer and Meyer (1972) call the "category-search model." It assumes that the subject searches through instances of the categories in memory seeking a match for the stimulus word. Such a model predicts that as the number of categories or words in the memory set increases, reaction time for the category-matching task should increase at a greater slope than reaction time for the word-matching task. This is because each additional target category adds more instances that must be searched, whereas each additional target word only adds one, the word itself. This result was essentially what Juola and Atkinson found.

The model they ascribed to Collins and Quillian (1970a) assumed that subjects perform the category-matching task by retrieving their stored category for the stimulus word and comparing this to the given categories to see if it matches one of them. This model would predict that the slope for the two tasks should be about the same, and the intercept for the category-matching task should be greater than for the word-matching task. Their results clearly reject this

model. Although attributed to Collins and Quillian, this model is quite different from Quillian's theory, because the semantic search in Quillian's (1966) theory is assumed to spread in parallel from both categories and instances. When the categories are given first, as in Juola and Atkinson's experiment, then activation would spread out from the categories before the instance even appeared.

In order to explain our interpretation of Juola and Atkinson's results, it is necessary to describe their procedure in more detail. They chose 10 large categories and 12 common instances from each category as stimuli. This makes a total of 120 instances in all. In the word-matching task, they presented from 1 to 4 of the 120 instances as targets on each trial. In the category-matching task, they presented from 1 to 4 of the 10 categories as targets on each trial. In both cases the negative stimuli were chosen from the same set of 120 instances. In the word-matching task, then, the discrimination necessary to categorize the stimulus was between one of the target words or a word that had not occurred as a target for a large number of trials (on the average about 24 trials earlier). In the category-matching task, however, the discrimination was between a word in one of the target categories and a word in one of the categories from a recent trial (on the average about 2 trials earlier). The discrimination, therefore, was rather easy in the word-matching task and quite difficult in the category-matching task.

What we think must be happening in the task is that the discrimination between positive and negative responses is made (at least partly) on the basis of activation level. In the category-matching task, as the number of categories increases, the amount of activation allocated to each category decreases (see Assumptions 2 and 3). Furthermore, activation will be left over from previous trials on the categories corresponding to negative instances, though it will have partly decayed (Assumption 3). For example, suppose "tree" was a target category on a particular trial and "body part" was not,

but "body part" was a target on a previous trial. Then "tree" will have a higher activation level than "body part," but the difference will not be very large because "body part" was presented so recently. If a positive instance such as "oak" is presented, it will intersect with "tree"; if a negative instance such as "arm" is presented, it will intersect with "body part." The less activation on "tree," which depends on how many other targets there are, the harder it is to discriminate that it is in fact a target or that "body part" is not a target. The more difficult the discrimination is, the longer it takes to make, and thus there will be a fairly large effect of the size of the target set on reaction time.

In the word-matching task, however, the difference in activation level as the number of targets is varied will not be so critical a factor. This is because the absolute difference between the activation level of targets and nontargets is so much greater in the word-matching task, given Juola and Atkinson's experimental procedure. That is to say, each nontarget was presented as a target so many trials previously (approximately 24) that the activation level for a nontarget would have decayed (Assumption 3) to a very low activation level as compared to any target. Hence, the large absolute difference in activation levels between targets and nontargets makes the differences due to target-set size relatively unimportant.

In conclusion, the spreading-activation theory's explanation of Juola and Atkinson's (1971) results is that the effect of differences in activation level due to target-set size matter more when the discrimination is difficult and matter less when the discrimination is easier. Furthermore, as Juola and Atkinson point out, there are two aspects of their data (namely, the fact that the data for positive responses are not linear, and the marked recency effects in the serial position curves) that fit much better with a parallel model, such as spreading-activation theory, than they do with a serial model, such as the one they derive from Landauer and Freedman (1968).

There are two implications of this view that could be tested fairly easily. One is that reaction time to decide that an instance such as "arm" is a negative instance will depend on the recency with which "body part" was presented as a target category. The more recent its presentation, the longer it will take to say "no." A more global implication is simply that the slope of the curve with respect to target size depends on those factors that affect the difficulty of discrimination, such as the recency with which negative instances occurred as targets (and probably as nontargets as well). This has important implications for the memory-search literature as a whole.

Conrad's Study

Using a true-false reaction time technique for sentences (e.g., the task is to decide whether "A salmon can eat" is true or false), Conrad (1972) found results which she interpreted as contradictory to Quillian's (1966, 1969) theory of semantic processing. In fact, the results of her study are quite close to what Quillian's theory would predict given Conrad's methodology.

In her first experiment, which was like the Collins and Quillian (1969) study, Conrad selected 2-level and 3-level hierarchies from the common culture (e.g., salmon → fish → animal) and properties associated with the objects at different levels. Then she constructed sentences with instances, such as "salmon," from the lowest level and properties from all three levels. The results Collins and Quillian found were that reaction time increased as the property was farther removed from the instance in the hierarchy. The reason for the increases in reaction time according to spreading-activation theory is that as the instance and property are farther apart in the hierarchy, it takes activation longer to spread between them and to trigger an intersection (and perhaps to evaluate the path found as well).

Unlike Collins and Quillian (1969), Conrad (1972) broke down the properties in her sentences into three groups on the basis of the frequency (high, medium, low) with which people generated each property, given

the different objects in the hierarchies. Another difference from Collins and Quillian's study is that she collected data over 5 days by repeating all the sentences each day.

The results of her first experiment were generally in the same direction as the increases in reaction time that Collins and Quillian found when the property was farther removed from the instance in the hierarchy. There was one reversal in her data out of nine comparisons, and this occurred for the high-frequency properties, where it was not unexpected given the weak theory of cognitive economy (as we will argue below). However, the increases she found were much smaller on the average than those of Collins and Quillian. The weak theory of cognitive economy predicts that people store a property with whatever instance it is linked to in a sentence, so Conrad's repetition of sentences over 5 days should lead to the smaller reaction time increases she found. This is because an inference necessary on the first day would be less likely on the second day, and so on. Conrad, in fact, reports a large Level $\times$ Day interaction.

In general, Conrad found that the higher the frequency of the property, the smaller the increases between levels. Given the weak theory of cognitive economy, we would expect that high-frequency properties are more likely to be stored at several levels in the hierarchy, because they are more likely to be encountered in contexts involving specific instances. For example, "leaves" are more likely to be stored as a property with particular types of trees (such as "maple" and "oak") than is "bark," because leaves are a higher frequency property. Thus, the effect of property frequency found by Conrad is consistent with the weak theory.

Conrad (1972) argued that Collins and Quillian's (1969) results could be explained by a confounding of property level and property frequency. Her argument was that the sentences Collins and Quillian used may have been based on high-frequency properties for Level 1 sentences (e.g., "A salmon is pink"), moderate-frequency properties for Level 2 sentences (e.g., "A salmon has fins"), and low-frequency properties for

Level 3 sentences (e.g., "A salmon can eat"). To support her argument, she showed that if one plots her reaction time data in the above way, one obtains approximately the same slope as Collins and Quillian did, whereas if one plots the slope for low-, medium-, or high-frequency properties separately, one obtains much smaller slopes.

However, there are two weaknesses in Conrad's argument. First, she did not use her frequency data to evaluate systematically the frequency of the properties in Collins and Quillian's sentences, so her conjecture about such a confounding had no empirical basis. Collins and Quillian (1969) did obtain subject ratings of importance of the property for the relevant level concept, which should correlate with Conrad's frequency measure. These ratings averaged 1.90 for Level 1 sentences, 1.92 for Level 2 sentences, and 2.16 for Level 3 sentences (based on a 5-point scale, where 1 = very important and 5 = not important). These are small differences and certainly do not support the notion that the slope between Level 1 and Level 2 sentences was due to the confounding Conrad hypothesized. The difference between Level 3 sentences and the others may have contributed to the greater slope that Collins and Quillian found, but even that is doubtful. For those subjects who had sentences with all 3 levels, the slope was actually larger (approximately 100 msec rather than 75 msec), but this was offset by a group of subjects who were slower overall and saw only Level 1 and 2 sentences. So the latter group acted to cancel out any exaggeration of the slope due to the lower importance of Level 3 properties.

Second, the comparison Conrad (1972) made in plotting her data against Collins and Quillian's data compared data based on five responses to the same sentence with data based on one response to a sentence. As indicated above, the weak theory of cognitive economy predicts that repetition of a sentence makes an inference less likely and should reduce the slopes in the way Conrad found. A fairer comparison would be between her data on the first day and Collins

and Quillian's data. But even that comparison has the problem that she may well have included sentences of the kind, "A maple has leaves," where the property is a general property of trees, but where most people would store it as a property of maples as well. This suggests that the fairest comparison is between Collins and Quillian's data and her data on the first day for low-frequency properties, where the properties were least likely to be stored at more than one level. But we cannot make this comparison because she did not break down her data by days. In conclusion, the differences between the two experiments and the fact that the only relevant data do not particularly support the conjecture about a confounding of property level and property frequency make Conrad's argument rather tenuous.

It was Conrad's second experiment that appears more damaging to Quillian's theory, but here she made a crucial methodological change. She presented the object 1 sec before the property, and this turned the experiment into a priming study. In the study she presented properties true of the highest level nodes, together with objects (e.g., "salmon," "fish," or "animal") at different levels in the hierarchy. Therefore, she predicted from Quillian's theory that the lower level objects, such as "salmon," would take longer to confirm, since it would take activation longer to spread between lower level objects and higher level properties. But by presenting the object 1 sec before the properties and by using only high-level properties, she made it possible for her subjects to prepare during the interval by priming the object's superordinates. For example, if a subject saw "salmon," his best strategy was to retrieve the superordinates, "fish" and "animal," because the property to appear would be a high-level property, such as "eating." In these circumstances, there is little reason to expect systematic differences between objects such as "salmon," "fish," and "animal." Thus, this particular experiment had real methodological problems as a test of Quillian's theory, and it is weaker evidence *against* spreading-activation

theory than her first experiment is evidence for the theory.

Effects of Typicality and Semantic Relatedness in Categorization Tasks

In recent experiments, Rips et al. (1973), Rosch (1973), and Smith et al. (1974) have shown that reaction time in a categorization task corresponds very closely to ratings of how typical the instance is of the category. For example, robins and sparrows are considered typical birds whereas chickens and geese are not. The effect of typicality on reaction time is quite large even when frequency of the particular instances in the language is controlled. Like Smith et al., we would argue that the typicality effect is one more manifestation of the fact that semantic similarity speeds up positive decisions and slows down negative decisions. Such an effect has been found repeatedly (Collins & Quillian, 1969, 1970a, 1972b; Schaeffer & Wallace, 1969, 1970; Wilkins, 1971). While Landauer and Meyer (1972) argued that the evidence for similarity effects at that time was either questionable or artifactual, the evidence now seems so overwhelming that any viable theory must account for them. They are very damaging to the category-search model.

There are two reasons why spreading-activation theory predicts that atypical instances will take longer to categorize than typical instances. The most important reason derives from the way evidence is aggregated (see Assumptions 8-13). Because different connections that are found are combined as evidence, distinguishing properties can slow down a positive decision based on a superordinate connection or on matching properties. For example, the decision that a chicken is a bird (i.e., an atypical instance) might be made on the basis of a superordinate connection from "chicken" to "bird," which people learn because chickens are frequently referred to as birds. But the fact that people eat chickens, that they are raised on farms, and that they are rather large are all properties that distinguish chickens from most birds. If these distinguishing properties are found during the

memory search, as some are likely to be, they act to slow down the positive decision, because they are negative evidence. Similarly, matching properties can slow down a negative decision. For example, the decision that a goose is not a duck might be made on the basis of the difference in their necks (a distinguishing property) or simply because they are stored as mutually exclusive kinds of birds (see Assumption 12), but the matching properties that are found (e.g., their affinity to ponds, their webbed feet, their large size) will slow down the decision that they are different. The argument here is similar to that of Smith et al. (1974), which we will discuss in comparing the two theories.

The second reason for the typicality effects relates to those cases where a superordinate connection is found. As we indicated earlier, superordinate links differ in accessibility (or strength), and accessibility depends on use. If a person frequently uses the link that a robin is a bird, and less frequently uses the link that a chicken is a bird (assuming approximately equal frequency for chickens and robins), then the accessibility of "bird" from "robin" will be greater than from "chicken." Because of this, accessibility will be highly correlated with typicality ratings. All the factors acting to make a chicken or a goose an atypical bird in the real world will also act to make the use of the superordinate link in a person's mind from "chicken" or "goose" to "bird" infrequent. It is because they are atypical that the superordinate link is weak, and this will also act to slow down reaction time in making categorization judgments about atypical instances.

The way evidence is aggregated in the theory also explains the common finding (Collins & Quillian, 1970a, 1972b; Holyoak & Glass, 1975; Rips et al., 1974) that people are fast to decide that semantically unrelated concepts are different (e.g., that a book is not a dog). In comparing such concepts, there are not likely to be any superordinate connections, and almost all property connections will involve distinguishing rather than matching properties. Therefore, almost

all the connections found will constitute negative evidence, and subjects will be quite fast to reach the negative criterion in such cases. This too is similar to the explanation in the Smith et al (1974) model.

Recently, Holyoak and Glass (1975) have isolated two different cases where semantic relatedness or typicality does not produce the usual effect on reaction time for negative judgments. One case arises when the decision depends on what they call a contradiction and what we have called mutually exclusive subordinates. The other case arises when the decision depends on a counterexample.

In the first case, Holyoak and Glass found that people are faster to reject sentences such as "All fruits are vegetables" or "Some chairs are tables" than sentences such as "All fruits are flowers" or "Some chairs are beds." In these four sentences the two nouns are mutually exclusive subordinates. The difference between the sentences is that "vegetables" and "tables" are generated with high frequency, while "flowers" and "beds" are generated with low frequency, when subjects are given the frame "All fruits are . . ." or "Some chairs are . . ." and asked to produce a false sentence. This difference is in the opposite direction of the usual finding that negative judgments are slower when the two concepts are more closely related semantically. The explanation for this reversal according to the theory (and to Holyoak and Glass) is that people make these decisions not on the basis of distinguishing properties (though some might be considered), but because they are stored as mutually exclusive subordinates (Assumption 12). Generation frequency in this case is a measure of the strength of the connection between the two concepts and therefore of how long it will take to find the contradiction between the two mutually exclusive concepts.

The second finding of Holyoak and Glass (1975) involves sentences where people reject the sentence by finding a counterexample (Assumption 13). For example, "All animals are birds," can be rejected by finding another kind of animal, such as a mammal. In this case Holyoak and Glass varied

the production frequency of the predicate noun (e.g., "birds") independently of the production frequency of the counterexample (e.g., "mammals"). Their finding was that reaction time depended not on the production frequency of the predicate noun (which is a measure of the semantic relatedness of the concepts in the sentence) but on the frequency of producing a counterexample. Here again where a decision strategy that is not based on distinguishing properties is appropriate, the reaction time data do not depend on the semantic relatedness of the two concepts.

The importance of these two findings by Holyoak and Glass (1975), in our view, is that they demonstrate that different kinds of evidence can be involved in making categorization judgments. This suggests that approaches such as that of Meyer (1970) and Smith et al. (1974), which try to formulate a single strategy for making such judgments, will inevitably fail.

RELATION OF THE THEORY TO THE MODEL OF SMITH, SHOEN, AND RIPS

Quillian's (1966, 1969) theory was a forerunner of a number of global theories of semantic processing based on network representations, in particular those of Anderson and Bower (1973), Norman and Rumelhart (1975), and Schank (1972). These theorists have made important advances on the Quillian theory (especially in the representation of acts and causes) which in no way contradict the basic thrust of Quillian's theory. There are some differences between these theories and Quillian's, but the basic intent of this paper is to deal with those aspects of semantic processing where the model of Smith et al. (1974) is the major competitor to Quillian's theory.

Unlike the various network models, the model of Smith et al. represents concepts as bundles of semantic features. Their model has the virtues of being quite clear and explicit, and it agrees quite well with the reaction time data for categorization judgments, except for the Holyoak and Glass (1975) results. Because it is such an initially compelling model, we want to emphasize how it differs from spreading-activation

theory and point out what we think are its inherent difficulties.

In the model of Smith et al. (1974), the meaning of a concept is assumed to be represented by semantic features of two kinds: defining features and characteristic features. Defining features are those that an instance must have to be a member of the concept, and the model assumes that features can be more or less defining. Characteristic features are those that are commonly associated with the concept, but are not necessary for concept membership. For example, "wings" might be a defining feature of "birds" and "flying" a characteristic feature, since all birds have wings but not all fly. In a categorization task, the model assumes that the two concepts are first compared in Stage 1 with respect to all their features, both characteristic and defining. If the match is above a positive criterion, the subject answers "yes"; if it is below a negative criterion, the subject answers "no"; and if it is in-between, the subject makes a second comparison in Stage 2 based on just the defining features. If the instance has all the defining features of the category, the subject says "yes" and otherwise says "no." If the subject can decide in Stage 1, his reaction time will be faster than if he decides in Stage 2.

There are several minor differences between the model of Smith et al. (1974) and the spreading-activation theory that could be minimized by slightly changing their model. The difference in wording between comparing features in their model and finding links between properties in our theory is really a nondifference. But the distinction between defining and characteristic features has the inherent difficulty, pointed out through the ages, that there is no feature that is absolutely necessary for any category.⁴ For example, if one removes the wings from a bird, it does not stop being a bird. Furthermore, we doubt if people can make consistent de-

cisions as to whether a feature is defining or characteristic, either from time to time or from one person to another. Smith et al. recognized that features are more or less defining (or criterial), but they were forced into making the artificial distinction between defining and characteristic in order to have a two-stage model. Still, the model could be revised to work without the two stages and make essentially the same reaction time predictions.

The revision is as follows: If features are compared over time, as in Quillian's (1966) theory, then as the process goes on longer, more features will be compared (assuming features have different accessibilities). The comparison process can have a positive criterion and a negative criterion just as before, and features can be weighted by their criteriality. If the match at any point in time is above the positive criterion, the subject says "yes"; if the match falls below the negative criterion, the subject says "no"; and otherwise he goes on comparing features. Finally, if he is running out of relevant information, he says "I don't know." This is simply the Bayesian decision model described in Assumption 9 of the extended theory, where the evidence consists of matching and mismatching features as in the property comparison of Assumption 11.

Thus, we agree that a decision process similar to the one that Smith et al. (1974) postulate does occur for *some* categorization decisions. But there is a fundamental disagreement, because they argue that *all* categorizations judgments are made by comparing features of the instance and category, whereas we argue that people use whatever evidence they find, including superordinate links.

Because they exclude the use of superordinate links, the model of Smith et al. has several inherent difficulties. The most obvious is the assumption that even when people have superordinate information stored, they do not use it. While most people may not have learned some superordinate relations (e.g., that a beaver is a mammal, or a sled is a vehicle), there are many they have learned (e.g., that a wren is a bird, and a beaver is an animal). Why would they not

⁴ There is for living things a biologists' taxonomy, which categorizes objects using properties that are not always those most apparent to the layman. Thus, there are arbitrary, technical definitions that are different from the layman's ill-defined concepts, but this is not true in most domains. There is no technical definition of a game, a vehicle, or a country that is generally accepted.

use such information if it is stored? How in fact can they avoid using it? It is an unlikely model which postulates that people use information that is less relevant to make a decision, instead of information that is more relevant.

Another obvious difficulty with the Smith et al. (1974) model is that people seldom know the defining properties of concepts. For example, consider whether a whale is a mammal, a sponge is an animal, a bat is a bird, or a wren is a sparrow. In the Smith et al. model, these difficult (and slow) decisions would be made in Stage 2 on the basis of defining properties. But people generally have no idea what the defining properties of a mammal, an animal, a bird, or a sparrow are. Even if they know that one of the most criterial properties for being a mammal is that it bears its young alive, it seems highly unlikely that they know whether whales (or beavers for that matter) bear their young alive. Neither of the authors has any idea what properties of a sponge make it an animal, but if asked in an experiment whether a sponge was an animal, we would answer "yes," and we would be comparatively slow about it. The reason we would answer "yes" is simply that we were told at one time that a sponge is an animal. We were also told that a bat is not a bird, and if we had not been told, we fear we might have responded "yes" if asked whether a bat is a bird in a categorization experiment. The decision that a wren is not a sparrow would be made because they are mutually exclusive kinds of birds (See assumption 12). They are both small songbirds, and it is hard to believe that many people know what the defining features of a sparrow are that a wren does not have. The fact that there are cases where people must use superordinate information to make correct categorization judgments makes it unlikely that they do not use such information in other cases where they could make the decision simply by matching features or properties. This is one of the strongest arguments for a hybrid theory.

We would like to close this section by raising the question of why one should adopt such a complicated theory when the Smith

et al. (1974) model is simpler and predicts the reaction time data quite well. We have tried to stress the inherent difficulties that their model has in ignoring superordinate information and in relying on defining properties. Experimental tests can probably be devised that will show up those difficulties. We will suggest one such test, but first we might point out that the results of the Loftus (1973b) categorization experiment described earlier do not fit the Smith et al. model very well. If a person is merely comparing features between the instance and the category, then it should not matter whether the instance or category is presented first. It is the asymmetry in the superordinate connections that predicts the asymmetry Loftus found in reaction time, and it is hard to imagine how one could have an asymmetry of that kind in comparing features of two concepts.

One experiment that might show difficulties with the Smith et al. (1974) model is a categorization task. The categories and instances used are based on their multidimensional scaling of birds and animals on the one hand, and mammals and animals on the other. As both Collins and Quillian (Note 4) and Rips et al. (1973) report, subjects are faster at deciding that bird names are in the category "bird" than in the category "animal," whereas they are slower at deciding that mammal names are in the category "mammal" than in the category "animal." Collins and Quillian argue that this is the way people learn the superordinates: that pigeons are birds and lions are animals. Smith et al. argue that it is based on shared features, and they show by their scaling solution that most birds are closer to "bird" than to "animal," and most mammals are closer to "animal" than to "mammal." But there are several bird names that are closer to "animal" than to "bird" (in particular, "goose," "chicken," and "duck"; "pigeon" is equidistant), and there are several mammal names that are closer to "mammal" than "animal" (in particular, "deer," "bear," and "lion"; "horse" is equidistant). We would predict that even for those instances the above pattern would hold, whereas a pure feature-matching theory,

such as the Smith et al. model, makes the opposite prediction. So this is a possible test of the two theories. There are undoubtedly many other tests.

Finally, we want to explain why we have been led to such a complicated theory. In trying to write computer programs that answer different types of questions, it becomes apparent that any decision procedure that gives correct answers must be flexible enough to deal with many different configurations of knowledge in memory. This is because people have incomplete knowledge about the world (see Collins et al., 1975), and they often do not have stored particular superordinate links or criterial properties. Any realistic data base for a computer system will have this same kind of incomplete knowledge. Therefore, perhaps our strongest criticism of the Smith et al. (1974) model is that it breaks down when people lack knowledge about defining features.

While at one level this is a complicated theory, at another level it is a simpler theory than the Smith et al. model. By viewing superordinate links as highly criterial properties, the theory becomes a simple Bayesian model. It is only in specifying the particular configurations of knowledge that constitute positive or negative evidence for the Bayesian process that the theory becomes complicated. The difference between the two theories is that the Smith et al. model allows only one kind of evidence (matching or mismatching features), whereas the theory presented here allows other kinds of evidence as well. Thus the theory encompasses a revised version of the Smith et al. model as a special case of a more general procedure.

CONCLUSION

We have extended Quillian's spreading-activation theory of semantic processing in order to deal with a number of experiments that have been performed on semantic memory in recent years. The result is a fairly complicated theory with enough generality to apply to results from many different experimental paradigms. The theory can also be considered as a prescription for building human semantic processing in a computer, though at that level many details are omitted

about decision strategies for different judgments that arise in language processing (see Carbonell & Collins, 1973; Collins et al., 1975; Quillian, 1969). We would argue that the adequacy of a psychological theory should no longer be measured solely by its ability to predict experimental data. It is also important that a theory be sufficiently powerful to produce the behavior that it purports to explain.

REFERENCE NOTES

1. Wilkins, A. J. *Categories and the internal lexicon*. Paper presented at the meeting of the Experimental Psychology Society, Oxford, England, 1972.
2. Loftus, E. F. *How to catch a zebra in semantic memory*. Paper presented at the Minnesota Conference on Cognition, Knowledge, and Adaptation, Minneapolis, 1973.
3. Meyer, D. E., Schvaneveldt, R. W., & Ruddy, M. G. *Activation of lexical memory*. Paper presented at the meeting of the Psychonomic Society, St. Louis, 1972.
4. Collins, A. M., & Quillian, M. R. *Categories and subcategories in semantic memory*. Paper presented at the meeting of the Psychonomic Society, St. Louis, 1971.

REFERENCES

- Anderson, J. R., & Bower, G. H. *Human associative memory*. Washington, D.C.: V. H. Winston, 1973.
- Anderson, R. C., & McGaw, B. On the representation of meanings of general terms. *Journal of Experimental Psychology*, 1973, 101, 301-306.
- Bell, A., & Quillian, M. R. Capturing concepts in a semantic net. In E. L. Jacks (Ed.), *Associative information techniques*. New York: American Elsevier, 1971.
- Brown, R. W., & McNeill, D. The "tip of the tongue" phenomenon. *Journal of Verbal Learning and Verbal Behavior*, 1966, 6, 325-337.
- Carbonell, J. R., & Collins, A. M. Natural semantics in artificial intelligence. *Proceedings of the Third International Joint Conference on Artificial Intelligence*, 1973, 344-351.
- Collins, A. M., & Quillian, M. R. Retrieval time from semantic memory. *Journal of Verbal Learning and Verbal Behavior*, 1969, 8, 240-248.
- Collins, A. M., & Quillian, M. R. Does category size affect categorization time? *Journal of Verbal Learning and Verbal Behavior*, 1970, 9, 432-438. (a)
- Collins, A. M., & Quillian, M. R. Facilitating retrieval from semantic memory: The effect of repeating part of an inference. *Acta Psychologica*, 1970, 33, 304-314. (b)
- Collins, A. M., & Quillian, M. R. Experiments on semantic memory and language comprehension.

- In L. W. Gregg (Ed.), *Cognition in learning and memory*. New York: Wiley, 1972. (a)
- Collins, A. M., & Quillian, M. R. How to make a language user. In E. Tulving & W. Donaldson (Eds.), *Organization of memory*. New York: Academic Press, 1972. (b)
- Collins, A., Warnock, E. H., Aiello, N., & Miller, M. L. Reasoning from incomplete knowledge. In D. G. Bobrow & A. Collins (Eds.), *Representation and understanding: Studies in cognitive science*. New York: Academic Press, 1975.
- Conrad, C. Cognitive economy in semantic memory. *Journal of Experimental Psychology*, 1972, 92, 149-154.
- Fitts, P. M. Cognitive aspects of information processing: III. Set for speed versus accuracy. *Journal of Experimental Psychology*, 1966, 71, 849-857.
- Freedman, J. L., & Loftus, E. F. Retrieval of words from long-term memory. *Journal of Verbal Learning and Verbal Behavior*, 1971, 10, 107-115.
- Grober, E., & Loftus, E. F. Semantic memory: Searching for attributes versus searching for names. *Memory & Cognition*, 1974, 2, 413-416.
- Hollan, J. D. Features and semantic memory: Set-theoretic or network model? *Psychological Review*, 1975, 82, 154-155.
- Holyoak, K. J., & Glass, A. L. The role of contradictions and counterexamples in the rejection of false sentences. *Journal of Verbal Learning and Verbal Behavior*, 1975, 14, 215-239.
- Juola, J. F., & Atkinson, R. C. Memory scanning for words versus categories. *Journal of Verbal Learning and Verbal Behavior*, 1971, 10, 522-527.
- Landauer, T. K., & Freedman, J. L. Information retrieval from long-term memory: Category size and recognition time. *Journal of Verbal Learning and Verbal Behavior*, 1968, 7, 291-295.
- Landauer, T. K., & Meyer, D. E. Category size and semantic memory retrieval. *Journal of Verbal Learning and Verbal Behavior*, 1972, 11, 539-549.
- Loftus, E. F. Activation of semantic memory. *American Journal of Psychology*, 1973, 86, 331-337. (a)
- Loftus, E. F. Category dominance, instance dominance, and categorization time. *Journal of Experimental Psychology*, 1973, 97, 70-74. (b)
- Loftus, E. F., & Cole, W. Retrieving attribute and name information from semantic memory. *Journal of Experimental Psychology*, 1974, 102, 1116-1122.
- Loftus, E. F., Senders, J. W., & Turkeltaub, S. Retrieval of phonetically similar and dissimilar category members. *American Journal of Psychology*, 1974, 87, 57-64.
- Loftus, G. R., & Loftus, E. F. The influence of one memory retrieval on a subsequent memory retrieval. *Memory & Cognition*, 1974, 2, 467-471.
- Meyer, D. E. On the representation and retrieval of stored semantic information. *Cognitive Psychology*, 1970, 1, 242-300.
- Meyer, D. E. Correlated operations in searching stored semantic categories. *Journal of Experimental Psychology*, 1973, 99, 124-133.
- Meyer, D. E., & Schvaneveldt, R. W. Facilitation in recognizing pairs of words: Evidence of a dependence between retrieval operations. *Journal of Experimental Psychology*, 1971, 90, 227-234.
- Norman, D. A., & Rumelhart, D. E. *Explorations in cognition*. San Francisco: W. H. Freeman, 1975.
- Quillian, M. R. A revised design for an understanding machine. *Mechanical Translation*, 1962, 7, 17-29.
- Quillian, M. R. *Semantic memory*. Unpublished doctoral dissertation, Carnegie Institute of Technology, 1966. (Reprinted in part in M. Minsky [Ed.], *Semantic information processing*. Cambridge, Mass.: M.I.T. Press, 1968.)
- Quillian, M. R. Word concepts: A theory and simulation of some basic semantic capabilities. *Behavioral Science*, 1967, 12, 410-430.
- Quillian, M. R. The Teachable Language Comprehender: A simulation program and theory of language. *Communications of the ACM*, 1969, 12, 459-476.
- Rips, L. J., Shoben, E. J., & Smith, E. E. Semantic distance and the verification of semantic relations. *Journal of Verbal Learning and Verbal Behavior*, 1973, 12, 1-20.
- Rosch, E. On the internal structure of perception and semantic categories. In T. E. Moore (Ed.), *Cognitive development and acquisition of language*. New York: Academic Press, 1973.
- Schaeffer, B., & Wallace, R. Semantic similarity and the comparison of word meanings. *Journal of Experimental Psychology*, 1969, 82, 343-346.
- Schaeffer, B., & Wallace, R. The comparison of word meanings. *Journal of Experimental Psychology*, 1970, 86, 144-152.
- Schank, R. C. Conceptual dependency: A theory of natural language understanding. *Cognitive Psychology*, 1972, 3, 552-631.
- Schvaneveldt, R. W., & Meyer, D. E. Retrieval and comparison processes in semantic memory. In S. Kornblum (Ed.), *Attention and performance IV*. New York: Academic Press, 1973.
- Smith, E. E., Shoben, E. J., & Rips, L. J. Comparison processes in semantic memory. *Psychological Review*, 1974, 81, 214-241.
- Sternberg, S. High-speed scanning in human memory. *Science*, 1966, 153, 652-654.
- Stone, M. Models of choice-reaction time. *Psychometrika*, 1960, 25, 251-260.
- Wilkins, A. J. Conjoint frequency, category size, and categorization time. *Journal of Verbal Learning and Verbal Behavior*, 1971, 10, 382-385.
- Wittgenstein, L. *Philosophical investigations* (G. E. M. Anscombe, trans.). Oxford: Blackwell, 1953.

Subjective Duration and Psychophysics

Hannes Eisler

University of Stockholm, Stockholm, Sweden

Three models concerned with the behavior of subjects estimating two successive durations are proposed. The accepted model assumes that the subjective total duration (sum of first and second durations) and the second duration are each accumulated in a separate sensory register. In a duration-matching experiment, for instance, the difference between the contents of the two registers is matched to the content of the second register. This model is accepted because (a) it does not include any memory, thereby eliminating certain difficulties connected with coding and storing of duration, (b) it copes with characteristic features of duration discrimination, and (c) it can satisfactorily explain data obtained in four scaling experiments, namely, magnitude estimation, matching, halving, and doubling of 10 durations between 1.3 and 20 sec. As a by-product, Stevens' power law is uniquely derived, and exponents are computed from matching data, thus eliminating the subjects' numerical behavior. The model also accounts for the time-order error for time.

The present investigation can be seen as lying at the intersection of two problem complexes: the validity of psychophysical theory for direct scaling and the psychophysics of time perception. The connection between the two is obvious—as will be shown, it is precisely the peculiarities of the time sense that make it possible to validate direct psychophysical scaling methods, to demonstrate that Stevens' power law is the only possible psychophysical function, and to determine the exponent for duration without involving the subjects' numerical behavior.

Three models are proposed to describe the strategy applied by a subject (or, perhaps better, the process in a subject) when he is confronted with two successive time intervals and is required to deal with some relation between them, for example, by telling which was the longer or by adjusting the second to match the first. The models are independent of the choice of any particular subjective scale (psychophysical function),

but a number of nonquantitative arguments seem to favor one of them. In order to compare the models quantitatively as well as connect them with scaling theory, a series of experiments was carried out, namely, magnitude estimation (ME), equal setting (ES), halving (HS), and doubling (DS) of 10 durations varying between 1.3 and 20 sec. The typical outcome of ES for one subject is taken as a starting point to derive Stevens' power law and other psychophysical relations. Based on this theoretical development, the experimental results are dealt with and the three models are reexamined in the light of these findings.

A closer analysis of a subject's task when dealing with two successive durations brings out two problems specifically pertinent to the perception of time:

1. The experimenter is not really in control of the stimulus: Time continues to elapse after the end of the first duration, during the interstimulus interval (ISI), and during the second stimulus duration. We might thus ask whether a subject is capable of ignoring the steady flow of time and focusing on the limited interval presented while waiting until the response can be given.

2. According to the traditional view, the subject has to perceive the first duration and retain "it" in memory until the offset

This work was supported by the Swedish Council for Social Science Research. I am indebted to K. Szczygiel, who designed and built the equipment, and to S. Holm for mathematical, K. Forsgren for experimental, and M. Vellner for computational assistance.

Requests for reprints should be sent to Hannes Eisler, Department of Psychology, University of Stockholm, Box 6706, S-113 85 Stockholm, Sweden.

TABLE 1
THREE MODELS FOR PROCESSING SUCCESSIVE TIME INTERVALS

Model	Standard		Variable	
	In subjective terms		In physical terms	
A	ψ_s	ψ_v		
B	ψ_s	$\psi_T - \psi_s$	$\alpha(\Phi_s - \Phi_{0s})^\beta$	
C	$\psi_T - \psi_v$	ψ_v	$(\Phi_s - \Phi_{0s})^\beta$	$(\Phi_v - \Phi_{0v})^\beta$
			$\alpha(\Phi_s + \Phi_v - \Phi_0)^\beta - (\Phi_v - \Phi_{0v})^\beta$	$\alpha(\Phi_s + \Phi_v - \Phi_0)^\beta - (\Phi_s - \Phi_{0s})^\beta$

Note. The terms *standard* and *variable* in the heading refer to the subject's point of view; subscripts *s* and *v* to standard and variable from the experimenter's point of view. Subscript *T* denotes total: $\Phi_T = \Phi_s + \Phi_v$.

of the second duration, when he gives his response. If the second duration is long, the first duration will have to be coded and transferred to the long-term memory (LTM), a process that might require more time than allowed by the ISI. The three models that follow solve these two problems (1 and 2) in different ways or fail to treat them at all.

THREE MODELS FOR PROCESSING SUCCESSIVE TIME INTERVALS

The models are most easily understood in terms of an equal-setting task, in which the subject is required to adjust a duration (the variable) so that it subjectively matches a preceding duration (the standard). The terms *standard* and *variable* are used in the traditional sense, from the experimenter's point of view. The models differ in what constitutes the standard and the variable from the subject's point of view.

The models are summarized in Table 1, where ψ refers to subjective (perceived) time, and subscripts *s*, *v*, and *T* refer to standard, variable, and total time from the experimenter's point of view (total time = sum of standard and variable).

Model A is the traditional one, where experimenter and subject agree on standard and variable. This model makes no explicit provision for Problems 1 and 2.

Model B assumes agreement between experimenter and subject regarding the standard. The variable from the subject's point of view, however, is the difference between the total duration and the standard duration. The subject could argue that this difference equals the variable, and if parallel processing (in some sense) is assumed, he could go on

coding and storing in LTM the standard duration while the variable duration is in progress, thus also for short ISIs. Model B thus takes account of Problem 2, but the subject still has to attend to the standard. (Model B is implied by an argument from Ekman & Frankenhaeuser, Note 1.)

Model C assumes agreement between experimenter and subject regarding the variable, but the standard from the subject's point of view is the difference between the total and the variable duration. In the same manner as in Model B, the subject could argue that this difference equals the standard. Model C solves both Problems 1 and 2: The ongoing time is "used" until the subject's response, and no memory is required at all, again provided that some kind of parallel processing takes place.

Note that if time perception were veridical, that is if subjective time were proportional to physical time, the three models would elicit the same behavior—the matched durations would be the same (and equal to the standard). And the subject would be right in his "argument." For nonproportional psychophysical functions, on the other hand, the three models yield different outcomes.

Figure 1 gives the models as block diagrams. Though the boxes labeled LTM and Sensory Register denote purely hypothetical constructs, a few distinctive features of a long-term memory and a sensory register may be noted: (a) A sensory register can be cleared, whereby its content is lost. The content of a long-term memory is lasting. (b) The content of a sensory register is not coded except in the trivial sense that, for example, time cannot be represented as such.

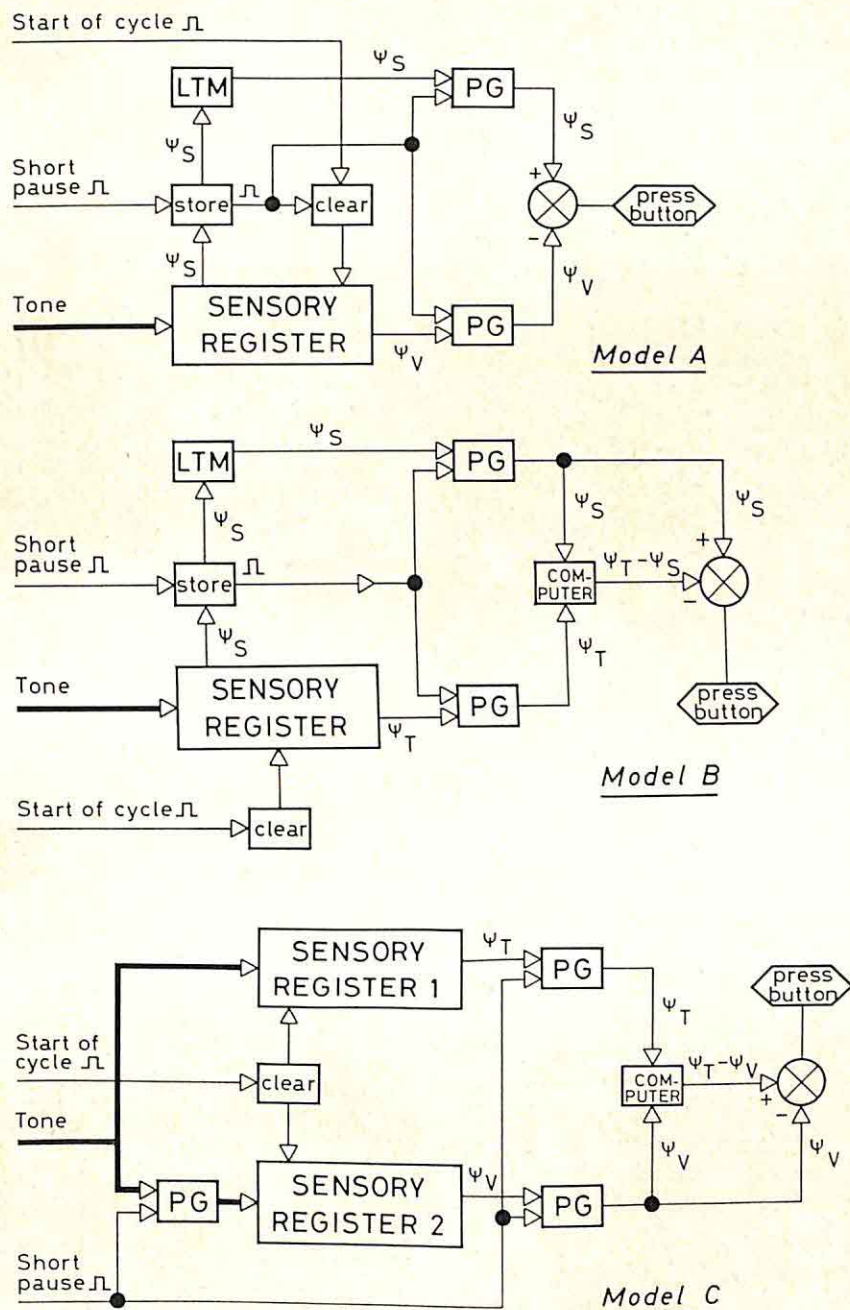


FIGURE 1. Block diagrams for the three models. PG = pulse-controlled gate (closed at start of cycle); LTM = long-term memory; $\otimes$ = comparator; Λ = a pulse.

The content of a long-term memory is coded more or less efficiently. (c) The content of a sensory register typically changes continuously, though not necessarily throughout the period of its use. The content of any memory is fixed.

The properties of a long-term memory and a sensory register, as listed above, are only sketched and call for qualifications, which, however, are not important in the present context. Unlike Point (c), Points (a) and (b) indicate features that a sensory register

might share with a short-term or an iconic memory.

The diagrams are drawn with the experiments mentioned in the introduction in mind. A "cycle" is a sequence of two durations, standard and variable, separated by a short pause, the ISI. The durations are "filled" and marked by a tone. The subject turns off the variable by pressing a button when it is experienced as equal to the standard (= zero output from the comparator). Thus there are two momentary events (start of cycle and ISI) and an event lasting for some time (the tone).

According to Model A, the sensory register is cleared at the start of the cycle. It then accumulates the duration indicated by the tone until the short pause, symbolized by a pulse. This pulse starts the coding and storing in the LTM of the content of the sensory register. When this has been effected, a new pulse resets the sensory register and opens the two pulse-controlled gates to the comparator. There the content of the LTM is continuously compared with the content of the sensory register. We see that if the time required for coding and storing exceeds the ISI, the first part of the variable duration will be lost, making it impossible for the subject to fulfill his task in the intended manner.

This is not the case in Model B, which differs from Model A in that the sensory register is not reset after the transfer of its content to the LTM. In the comparator the content of the LTM is continuously compared with the computer output, which is the difference between the total and the standard duration. The subject can carry out his task as intended except for cases where both the ISI and the variable are extremely short. Then he misses the right moment to press the button.

Whereas both Model A and Model B make use of the LTM, this is replaced by a second sensory register in Model C. Both sensory registers are cleared at the start of the cycle, and the first one accumulates the total time. The short pause pulse starts the second register, where the variable duration then accumulates, and opens the two gates. In the comparator the variable duration is

continuously compared with the computer output, which is the difference between the total and the variable duration.

The following analogy may make the processes more understandable. Imagine a subject equipped with three watches of identical appearance, Watch 1, Watch 2, and Watch 3. The dials are not circular but oblong (like certain car speedometers), they have no scale or mark except for a zero, and the hands move linearly. Watches 1 and 2 are stop-watches; Watch 3 does not have any works, but the hand is adjustable. In all cases two of the three watches are used, and they are placed with the zero points aligned.

According to Model A, the subject uses Watches 1 and 3. He starts Watch 1 at the start of the standard and stops it at the short pause. Then he moves the hand of Watch 3 to the position of the hand of Watch 1 (= stores the standard duration in LTM), restarts Watch 1 at the start of the variable duration and waits until the hand of Watch 1 coincides with the hand of Watch 3, whereupon he presses the button.

According to Model B, Watch 1 is started at the onset of the standard; at the short pause the hand of Watch 3 is set to the position, which the hand of Watch 1 had at that time (= the standard duration is stored in LTM). When the distance from the zero point to the hand of Watch 3 equals the distance between the two hands (one stationary, the other moving), the subject presses the button.

According to Model C, Watch 1 is started at the onset of the standard and Watch 2 at the onset of the variable. When the distance between the zero point and the hand of Watch 2 equals the distance between the hands (both moving), the subject presses the button.

It is easily realized that in Model A the subject would not have enough time to move the hand of Watch 3 and restart Watch 1 before the onset of the second duration, if the ISI is short. In Model B the time required to move the hand of Watch 3 might exceed the time of the ISI plus the second duration, if both are very short. None of these difficulties beset Model C.

Finally, let us note the following. If the movements of the hands are uniform, the outcome (moment of button press) is the same for all three models. This would correspond to a veridical perception of time, and the three models would be indistinguishable. Other than proportional psychophysical functions correspond to accelerated or retarded movements of the hands in the analogy, entailing different outcomes for the three models.

DURATION DISCRIMINATION AND THE THREE MODELS

Allan and Kristofferson (1974) list three findings that characterize duration discrimination:

1. A subject's ability to discriminate a constant difference in duration is independent of the duration values.
2. Discriminability between two durations is independent of the ISI.
3. Forced-choice performance is not superior to single-stimulus performance.

The first two statements need some qualification: The first statement seems to be valid only for practiced observers, and the second holds only for ISIs exceeding 500 msec; for shorter ISIs discriminability *deteriorates*.

We will examine these findings within the frame of reference of our three models. For the case of discrimination between two successive durations the subject's response depends on the output from the comparator at the offset of the second duration: +, -, or 0.

In continua other than time (as well as for unpracticed observers in duration discrimination), discriminability deteriorates with increasing stimulus magnitude for a given difference, particularly when the stimuli are presented in succession. This may be attributed to the increase in variability of subjective magnitude with magnitude, particularly for the memory of the first stimulus. Time perception according to Models A and B should behave in the same way. Model C, however, can explain why the length of the *duration* does not affect discriminability (Allan and Kristofferson's first finding). First, no memory is involved. Second, if we assume (a) that there is a correlation of

unity between the errors in the outputs from the two sensory registers—reasonable because of the simultaneity—and (b) that the errors are roughly proportional to the durations,¹ they become the same for both inputs to the comparator and thus cancel: $\psi_T + 2\epsilon - (\psi_V + \epsilon)$ yields an error of ϵ , that is, the same as in ψ_V . (Perhaps it may be concluded that a naive observer behaves in accordance with Models A or B and only a practiced one with Model C.)

The influence of the ISI found for continua other than time may be attributed to effects on the memory that progress with the length of the ISI and possibly to disturbances in the perceiving of the second stimulus while the first is being coded and stored. Since Model C does not require any memory, it eliminates this factor and thus conforms to Allan and Kristofferson's second finding. One would have to assume, however, that Sensory Register 1 is switched off during the ISI. The disruption of performance for ISIs < 500 msec might be due to the time required for switching Sensory Register 1 on and off.

Allan and Kristofferson's third finding is predicted by Model C under the assumption that the subject does not follow the single-stimulus instructions as intended by the experimenter. Even singly presented durations are given in a sequence, and rather than storing the standard in the LTM and comparing it with the current stimulus, the subject compares the current stimulus with the preceding one in the way suggested by Model C. If the output from the comparator is +, his answer is "short," if it is -, "long." After each response the register holding the total duration is cleared, and the tasks of the two registers are switched with each new stimulus, so that the new duration accumulates in the presently cleared register, whereas the other register continues to accumulate the new duration on top of the previous one. This implies that the subject is carrying out a forced-choice task even when instructed according to single stimulus, and that a single-stimulus experiment proper simply cannot be carried out with durations.

¹ Ekman's law (Ekman, 1956, 1959). See also Eisler (1962) and Teghtsoonian (1974).

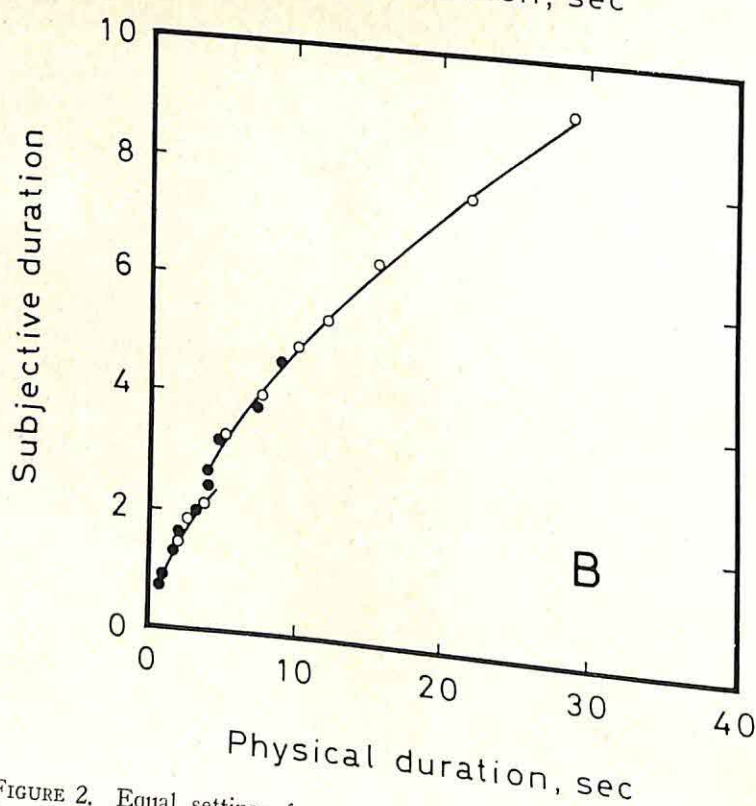
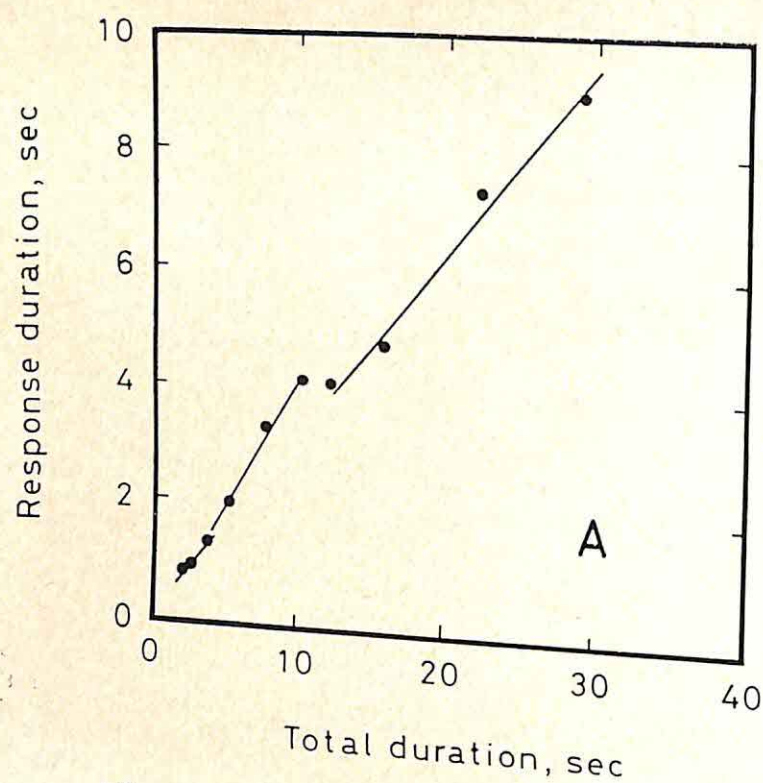


FIGURE 2. Equal settings for Subject 1. A: Response duration Φ_r as a function of total duration Φ . The lines are fitted by least squares. B: Subjective duration as a function of physical duration, showing the two segments of the psychophysical function. Open circles refer to standards, closed circles to variables. The ordinates of the points are computed; the abscissas are empirical values.

The advantage of the forced-choice method in always presenting two different durations in a trial is offset by the subject's opportunity of repeating his latest answer when the output from the comparator is 0. However, though the overall results of single stimulus and forced choice should be approximately the same according to Model C, a difference should be detectable in the microstructure of the data: In single stimulus both a correct and an incorrect judgment should show a tendency to perpetuate itself as long as the next stimulus equals the preceding one. All of these considerations favor Model C.

RATIO SCALING OF SUBJECTIVE DURATION: THEORY

Most of what is said in this section is pertinent also for continua other than time. Stevens' power law will be derived from precisely specified assumptions. Furthermore, it will be shown that combining Model C with the power law has two interesting consequences: (a) possible alternative psychophysical functions (cf. Treisman, 1964) have to be rejected, and (b) the exponent can be estimated from equal-setting data, that is, without involving the subject's numerical behavior.

The theoretical development is undertaken with the results of the ES experiment for Subject 1 as the starting point. In Figure 2A these data are given as a plot, where the variable duration is plotted against the total duration, in accordance with Model C. This particular plot has been chosen for demonstration because it gives the clearest picture of the main features of all 36 plots (3 experiments, 12 subjects) with an almost even distribution of the 10 data points among three straight lines. (As to fit, it ranks only 27th among the 36 sets of data.) The other plots are similar, though a number of them yield only two straight lines, and a few only one. (Four plots show slight curvilinearity.) Two features are apparent in Figure 2A: The plot is segmented, in that it shows discontinuities, and the segments are linear.

The Psychophysical Function

The linearity observed does not depend on the choice of model. It is easily shown that a plot of variable against standard duration (Model A) or of standard duration against total duration (Model B) likewise is linear.

The derivation given below is based on two definitions, two assumptions, and one empirical finding. It will be shown later that one of the assumptions follows from Model C; the other assumption is confirmed by the experimental results.

The empirical linearity of the data can be described by²

$$\Phi_V = a\Phi + b \quad (1)$$

(empirical finding), where Φ denotes the physical value of the standard duration (in general, here of the total duration), and Φ_V the variable duration. Furthermore, we have

$$\psi = f(\Phi) \quad (2)$$

(first definition), and

$$\psi_V = g(\Phi_V) \quad (3)$$

(second definition), where ψ denotes subjective values and f and g the psychophysical functions for standard and variable durations, respectively. If the subject carried out an r setting ($r = 1, .5$, and 2 for ES, HS, and DS, respectively), we have

$$\psi_V = r\psi \quad (4)$$

(first assumption). Equations 1-4 yield

$$rf(\Phi) = g(a\Phi + b). \quad (5)$$

The functional Equation 5 is too general to allow a solution. We can pick f at will and can always find the corresponding g . However, a simple restriction makes Equation 5 uniquely solvable.

The second assumption is that the two psychophysical functions, f and g , are independent of the particular task, that is, independent of r . In other words, the psychophysical functions should be the same, regardless of what ratios are to be set by the subjects. This requirement is an obvious prerequisite for any psychophysical theory of some generality. This invariance assumption

² A more detailed derivation of the psychophysical function will be found in Eisler (1974).

tion entails that a variation in r is reflected only by a variation in a and possibly b .

Taking the derivative of Equation 5 with respect to r yields

$$f(\Phi) = (a'\Phi + b') \cdot g'(a\Phi + b), \quad (6)$$

and with respect to Φ yields

$$r \cdot f'(\Phi) = a \cdot g'(a\Phi + b). \quad (7)$$

Dividing Equation 7 by Equation 6 yields

$$\frac{r \cdot f'(\Phi)}{f(\Phi)} = \frac{a}{a'\Phi + b'}, \quad (8)$$

and integrating Equation 8 with respect to Φ yields

$$r \cdot \ln f(\Phi) = \frac{a}{a'} \cdot \ln |a'\Phi + b'| + C_1(r),$$

or

$$f(\Phi) = C(r) \cdot (a'\Phi + b')^{a/a'}. \quad (9)$$

Remembering that $f(\Phi)$ is independent of r , $f(\Phi)$ is rewritten in the following way:

$$\psi = f(\Phi) = \alpha(\Phi - \Phi_0)^\beta, \quad \Phi > \Phi_0, \quad (10)$$

where α constitutes the scale unit, and Φ_0 the subjective zero. It is easy to show that the psychophysical function for the variable, g , is likewise a power function, featuring the same exponent but possibly different α and Φ_0 :

$$\psi_V = \alpha'(\Phi_V - \Phi'_0)^\beta, \quad \Phi_V > \Phi'_0. \quad (11)$$

Note that without any a priori assumption that the psychophysical functions for standard and variable are the same, we found this to be the case (with the possible exception of two parameters of secondary interest), and that the psychophysical function derived in this manner is Stevens' power law.³

It has already been shown (Ekman, 1958; Eisler, Note 2) that a power function is compatible with a linear $\Phi_V - \Phi$ relation; what is shown here is that a power function constitutes a unique solution to Equation 5.

³ Note that this derivation refutes the formulation of the power law as $\psi = \alpha\Phi^\beta + \psi_0$ (cf. the discussion in Galanter & Messick, 1961).

Discontinuous Psychophysical Functions

Consider Figure 2B, where subjective duration is plotted against physical duration. We see that the psychophysical function consists of two segments, that it exhibits a discontinuity, or break. Though breaks in psychophysical functions, particularly for the continuum of duration, have been found by quite a number of researchers (Eisler & Ekman, 1959; Reese, 1943; Ross & Katchmar, 1951; Taves, 1941; Ekman & Frankenhaeuser, Note 1), this issue has been largely neglected until now. The most convincing evidence as to the reality of these breaks is given by Richards (1964, 1973). At certain stimulus values his subjects produced bimodal response distributions, corresponding to the two functions around the break.

Both ψ and ψ_V values for the lowermost three points in Figure 2B lie on the lower segment, and ψ and ψ_V values for the uppermost four points lie on the upper segment, while for Points 4-6, ψ lies on the upper and ψ_V on the lower segment. This partition of the points into three groups is reflected in Figure 2A, where Φ_V is plotted against Φ : the lowest and the uppermost points lie on the same segment; the middle straight line, differing in slope from the other two, corresponds to the transition region, where the standards lie on the upper segment and the variables on the lower.

Thus, one break in the psychophysical function (Figure 2B) results in two breaks in the response-stimulus plot (Figure 2A). A detailed analysis of breaks in psychophysical functions can be found in Eisler (Note 2). Note that the break of the data in Figure 2 is at 4 sec, entailing a break at 4 sec on the x-axis and another at 4 sec on the y-axis of Figure 2A. Note further that the subject is not in "error" when he makes the same adjustment for different stimuli (Points 6 and 7); this is a simple result of there being two segments that differ in at least one parameter of the psychophysical function.

The parameters of the psychophysical function are obtained in the following way. Inserting Equations 10 and 11 into Equations

tion 4 we obtain

$$r\alpha(\Phi - \Phi_0)^\beta = \alpha'(\Phi_V - \Phi'_0)^\beta.$$

Rearranging yields

$$\Phi_V = \left(\frac{r\alpha}{\alpha'}\right)^{1/\beta} \cdot \Phi - \left(\frac{r\alpha}{\alpha'}\right)^{1/\beta} \cdot \Phi_0 + \Phi'_0. \quad (12)$$

Equating the coefficients of Equations 1 and 12 we obtain (see Ekman, 1958; Eisler, Note 2)

$$\left(\frac{r\alpha}{\alpha'}\right)^{1/\beta} = a,$$

that is,

$$\beta = \log(r\alpha/\alpha')/\log a, \quad (13)$$

and

$$-\left(\frac{r\alpha}{\alpha'}\right)^{1/\beta} \cdot \Phi_0 + \Phi'_0 = b. \quad (14)$$

Since the slope of the outer lines depends only on r and β ($\alpha' = \alpha$), they must be parallel. For the case shown in Figure 2, moreover, their intercepts are the same, implying also the same subjective zero. This happens often but not always. The slope of the middle line depends on r , β , and on the ratio α/α' of the scale units.

Denoting the slope of the outer lines by a , we obtain

$$\beta = \log r/\log a. \quad (15)$$

Denoting the slope of the middle line by a' , the ratio of the scale units is obtained:

$$\alpha/\alpha' = (a'/a)^\beta. \quad (16)$$

The intercepts and slope of the outer lines yield the subjective zero(s):

$$\Phi_0 = b/(1 - a), \quad (17)$$

where b is the intercept of the lowest line when it is the subjective zero of the lower segment of the psychophysical function that is sought, and of the uppermost line when it is the upper segment. Equation 14 provides a check. Denoting the intercept of the middle line by b' , the subjective zero of the upper segment by Φ_0 and of the lower segment by Φ'_0 , we have

$$-a'\Phi_0 + \Phi'_0 = b'. \quad (18)$$

When the data show only two straight lines or the third line consists of only one or two points, Equation 18 may be used to compute Φ_0 or Φ'_0 as the case may be.

Ratio Scales of Duration and the Three Models

What has been said up to now about scales is valid whichever model one chooses. Table 1 describes the three models also in terms of the psychophysical function. Because of the arbitrariness of the scale unit, it has been set to unity. If there are two segments with different scale units, the scale unit of the lower segment is set to 1, and that of the higher segment to α . Thus α may or may not equal unity in Table 1.

A remarkable feature of Models B and C is that the exponent for subjective duration can be computed from equal setting, that is, without any use of numbers in the experiment. For Model C, the constant r in Equation 5 becomes 1/2 (the variable subjective duration becomes half the total subjective duration). This constant of 1/2 belongs to the *scientific meta language* (Bergmann & Spence, 1944) from an epistemological point of view, since it emerged from the model, and not to the *object language*, as, for example, the constant .5 in a common half-setting experiment.

This fact proves that there is no alternative to the power function for the experiments described here by confirming the first assumption made in the derivation of Stevens' power law. What amounts to halving in the meta language is equal setting in the object language.

To recapitulate, the premises leading to the power function are as follows:

1. The ratio settings are linearly related empirically to the standards.

2. Model C (or B) is accepted, excluding possible other formulations of the subjects' behavior such as Equation 5, at least for equal setting.

3. The form of the psychophysical function is assumed to be independent of the prescribed ratio.

The first premise yields Equation 5, the uniqueness of which is guaranteed by the second premise, and the third premise sup-

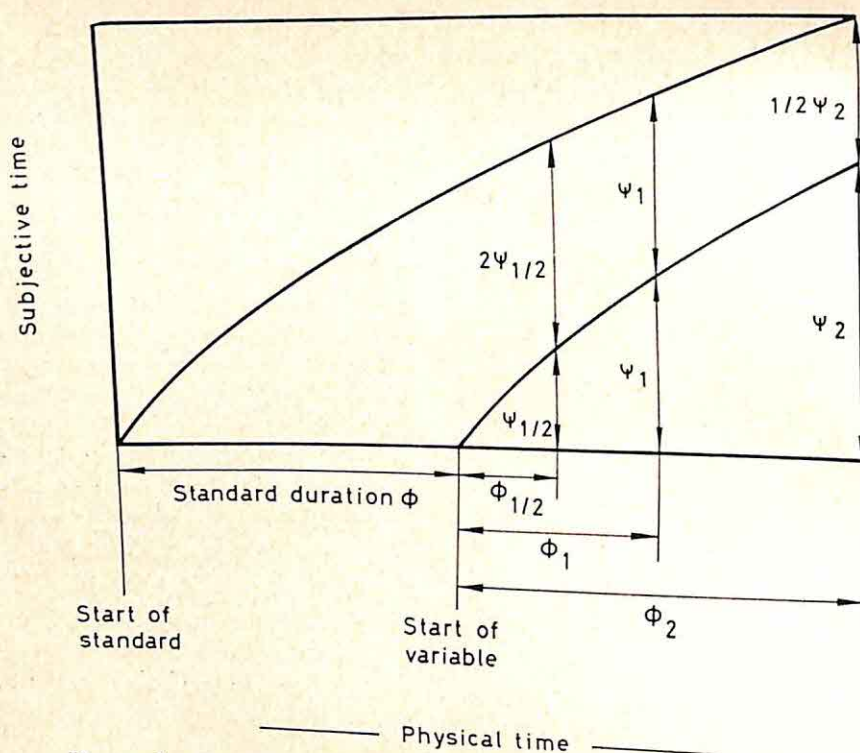


FIGURE 3. The ratio-setting tasks described in terms of the psychophysical function according to Model C. Subscripts refer to half, equal, and double setting.

plies the restriction that makes Equation 5 uniquely solvable.

In a later section the three models are examined in the light of the experimental results, and it is shown that Model C is to be preferred. Figure 3 describes the three ratio-setting tasks in terms of the psychophysical function for Model C.

EXPERIMENTS

Experimental Design

Using the same 10 time intervals as stimuli or standards, five experiments were carried out: reaction time (RT), magnitude estimation (ME), equal setting (ES), half setting (HS), and double setting (DS). The intervals (durations) were 1.3, 1.8, 2.5, 3.3, 4.5, 6.0, 8.1, 11.0, 14.8, and 20.0 sec. Thus they ranged from 1.3 to 20.0 sec and were equally spaced in logarithmic units. They were presented acoustically, as a strongly distorted tone from a small loudspeaker.

Each experiment consisted of a number of cycles in immediate succession. Every cycle was composed of four parts (see Figure 4): (a) a long pause (silent interval), followed by (b) one of the standard durations, followed by (c) a short pause (silent interval), that is, the ISI, followed by (d) the variable duration, indicated by the

same tone as the standard duration, and terminated by the subject's pressing a button. After the button press (which switched off the loudspeaker) the next cycle started immediately with the long pause.

In RT and ME the loudspeaker was not switched on in the fourth part of the cycle. The long pause amounted to 2 sec in all experiments, the ISI to 300 msec in all experiments except RT, where it was only 10 msec. (For technical reasons it could not be set to 0, but all reaction times existed only for the apparatus; it could not be perceived by the subject, since there was no tone during the fourth part of the cycle. It should be noted that the transition from standard to variable in the three ratio-setting experiments was thus marked by the short pause, that is, by a brief interruption of the tone.)

The 10 standard durations were given in pseudorandom order, randomized within a block of 30 times per block, so that each duration appeared three times per block. Then the same sequence of durations was repeated until a specified number of cycles had been attained.

Apparatus

The two parts of the apparatus were placed in different rooms and connected by cables. The sub-

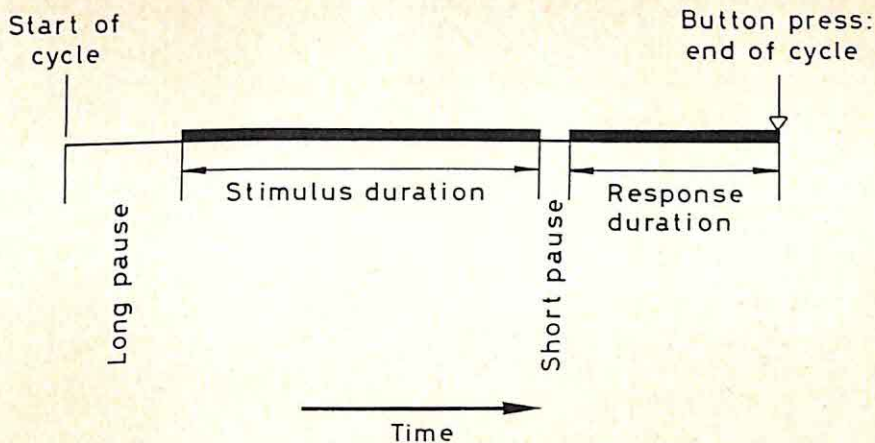


FIGURE 4. A cycle. For explanation see text.

ject's part consisted of a metal box with a small red lamp and the loudspeaker. Attached to this box by a cable was a grip with a pushbutton that activated a microswitch when pressed. This part of the equipment was monitored by the experimenter's part, which controlled the cycles, including the presentation of the standard durations and the registration of the response. The two pauses could be adjusted by two knobs and remained the same during any one experiment. The standard durations (30, in pseudorandom sequence of 10 different durations, see above) were given by means of a punched mylar tape in a closed loop. A cycle started with the reading of the next standard time and was terminated by the button press, which also initiated the next cycle. Standard time and variable time (response) were punched on paper tape (resolution: 10 msec) for later treatment with a high-speed computer. An experiment was stopped automatically after a predetermined number of cycles. The red lamp on the subject's box was switched off for the duration of the experiment, thus indicating to the subject the beginning and end of the experiment.

Procedure and Subjects

Twelve undergraduate students of psychology served as subjects as part of their course requirement.⁴ All subjects took part in all five experiments, which were carried out in a single session that lasted about 1½ hours including instruction time.

The end of an experiment was indicated to the subject by lighting the red lamp on his box. The experiments were carried out in the order of RT—ME—ES—HS—DS for all but six subjects, for whom the order of the last two experiments was reversed, so that DS preceded HS.

In RT the subjects were presented with 40 cy-

cles, in each of the other four experiments with 70 cycles. Since the number of cycles for each experiment was not a multiple of 30 (the number of cycles in a block), each experiment started, both intra- and interindividually, at a different position in the block.

The main purpose of the reaction-time experiment was to familiarize the subjects with the stimulus durations used, and it is therefore not considered further. The reaction times varied between 200 and 600 msec, most values being not far from 300 msec.

The magnitude-estimation experiment was carried out with a free standard. The instruction stressed that the subjects were to rely on their spontaneous experience. They should not "think in terms of seconds," nor should they count. Each estimate was to be entered in a mimeographed table, after which the subject was to press the button to obtain the next stimulus duration.

In the ratio-setting experiments the subjects were told that they had to listen to two durations, separated by a short pause. Pressing the button terminated the second duration, which was to be of the same length as (half the length of, twice the length of) the first one. Again, it was stressed that the task should be carried out in terms of spontaneous experience.

Results

In all four experiments the first 10 cycles were discarded as preliminary trials. Table 2 gives the means of the remaining 60 cycles (see exceptions below) for all 12 subjects individually.

Magnitude estimation. Every value in Table 2 is the arithmetic mean of six estimates. The standard deviations varied between 0 and 60% of the mean estimates, with 30% as a rough average.

Ratio settings. Of the 2160 adjusted durations, 44 were discarded as outliers, pro-

⁴ This number is usually sufficient to obtain group data of acceptable reliability. A closer examination of the data revealed, however, that the subjects behaved vastly differently, forcing me to treat the data individually.

duced by anticipatory error or slackened attention. The rule of thumb in discarding values was that when a value was more than twice or less than half the nearest value in the series of six, it was discarded. (In two cases two values were discarded from six.) Thus of the 360 response durations given in Table 2, 318 are based on six adjustments, 40 are based on five, and 2 on four. Though the number of outliers varied between subjects, all 12 subjects produced at least one each. The standard deviations varied between .05 and 7 sec, with typical values of less than 1 sec for the shortest durations, and 1-3 sec for the longest.

DATA ANALYSES

Magnitude-Estimation Experiment

The psychophysical function for the ratio settings has been shown to be a power func-

tion. Assuming that it is the same when the subjects are required to give magnitude estimates, linear plots should be obtained in log-log coordinates, when magnitude estimates are plotted against stimulus durations, provided the physical magnitude is measured from the subjective zero Φ_0 . These straight lines can be fitted with two parameters for $\Phi_0 = 0$, otherwise with three. However, the majority of the empirical curves exhibited one, or possibly more breaks, each break increasing the number of parameters by one or two, depending on whether the scale unit, the subjective zero, or both were different for the different segments.

A number of runs with a computer program that included Chandler's (1969) STEPIT compared the fit for different parameter combinations. As a result I decided to set $\Phi_0 = 0$ for all subjects, and also to accept

TABLE 2
RESPONSES FOR THE 12 SUBJECTS INDIVIDUALLY, WITH ARITHMETIC MEANS OF MAGNITUDE ESTIMATIONS
AND ARITHMETIC MEANS OF EQUAL, HALF, AND DOUBLE SETTINGS IN SECONDS

Subject	Standard duration (seconds)									
	1.3	1.8	2.5	3.3	4.5	6.0	8.1	11.0	14.8	20.0
Magnitude estimations										
1	2.00	2.00	3.33	6.33	6.67	8.67	8.33	12.33	12.33	18.00
2	3.00	4.17	6.00	7.00	8.17	10.50	13.17	16.33	23.17	30.00
3	1.50	1.67	2.08	2.83	3.00	3.50	3.67	4.50	5.00	6.83
4	3.00	3.33	7.33	8.33	11.00	15.00	20.50	26.67	34.17	48.33
5	2.50	4.33	4.67	7.33	7.17	10.00	10.17	14.83	56.67	42.50
6	4.33	6.33	18.33	24.17	29.17	30.00	41.67	65.83	85.83	96.67
7	1.00	1.50	2.17	2.50	3.33	3.83	4.50	6.67	9.17	11.17
8	2.50	2.92	4.58	5.00	8.67	11.67	15.00	18.33	26.67	36.67
9	2.00	2.08	2.42	3.00	4.50	5.17	6.83	8.67	11.00	13.83
10	5.83	8.17	10.33	14.83	18.33	32.50	38.33	51.67	66.67	94.17
11	1.67	2.17	2.83	4.17	5.00	7.17	9.17	12.00	16.50	21.00
12	5.00	8.33	16.67	17.50	23.33	34.17	33.33	40.00	53.33	63.33
Equal settings (seconds)										
1	.81	.93	1.32	1.97	3.27	4.12	4.05	4.73	7.36	9.05
2	1.35	1.57	2.29	3.39	4.59	4.85	5.84	7.38	10.15	13.52
3	.93	1.20	1.49	1.93	2.69	3.92	4.23	7.25	9.94	12.46
4	.62	.95	1.57	1.85	2.69	4.02	4.38	6.76	7.76	10.20
5	1.15	1.85	2.39	2.82	3.78	4.82	5.76	8.39	8.33	10.48
6	.65	1.37	2.43	3.24	3.73	4.99	6.63	8.22	10.33	14.05
7	1.02	1.45	2.86	3.39	4.80	5.71	6.26	7.75	7.94	13.23
8	.75	1.06	1.87	3.04	3.86	5.63	7.29	10.48	14.09	19.86
9	1.06	1.47	2.07	2.56	3.50	3.72	5.09	5.92	6.57	7.65
10	.46	1.31	2.03	2.33	3.30	5.09	6.92	8.88	13.85	15.50
11	.64	1.04	1.38	2.26	3.14	5.20	5.13	9.09	10.67	12.32
12	1.06	1.44	2.16	2.90	4.25	4.71	5.68	6.87	8.36	10.83

TABLE 2—(Continued)

Subject	Standard duration (seconds)									
	1.3	1.8	2.5	3.3	4.5	6.0	8.1	11.0	14.8	20.0
Half settings (seconds)										
1	.46	.52	.71	1.12	1.41	1.99	2.20	3.08	3.85	5.41
2	.57	.69	1.16	1.52	2.16	3.35	3.84	4.59	7.02	9.41
3	.77	.62	1.05	1.37	2.03	2.75	3.77	6.43	6.38	12.90
4	.31	.47	.77	.96	1.44	1.94	2.38	3.10	4.87	6.70
5	.72	.54	1.01	1.13	1.44	2.59	2.66	3.25	3.89	6.09
6	.60	.86	1.65	2.42	2.74	4.45	4.89	6.56	8.43	10.01
7	.70	1.29	2.38	3.07	3.06	4.74	4.66	5.39	6.98	7.31
8	.50	.51	.86	1.07	1.93	2.30	3.52	4.25	7.45	9.35
9	.48	.65	1.11	1.15	1.99	2.30	2.96	3.73	4.12	4.06
10	.42	.74	1.08	1.87	2.64	3.76	5.44	7.60	7.08	8.98
11	.56	.58	.69	.76	1.20	1.89	2.41	3.74	5.25	5.67
12	.95	.92	1.63	2.72	2.34	3.26	4.45	5.80	7.18	8.10
Double settings (seconds)										
1	2.17	2.28	3.00	3.83	3.65	6.73	7.47	7.89	8.66	11.57
2	2.18	2.78	4.44	4.54	7.14	8.77	9.36	11.07	13.59	20.60
3	3.78	3.60	3.93	4.74	6.94	8.91	9.11	13.68	16.13	20.48
4	1.96	2.36	3.52	4.11	4.39	6.29	7.81	11.31	11.84	19.15
5	3.34	3.98	6.38	5.89	11.87	8.28	12.16	15.81	14.54	22.01
6	3.01	4.32	6.27	7.98	9.64	11.88	12.91	12.30	13.58	20.87
7	2.28	3.11	5.12	8.04	7.59	9.26	10.22	13.28	16.12	19.43
8	2.40	3.00	4.98	5.46	7.89	10.14	16.42	20.76	23.89	34.68
9	2.04	2.66	3.38	4.62	6.67	7.97	8.72	10.25	11.15	16.51
10	1.57	3.04	5.06	5.85	8.11	10.88	13.89	19.33	28.29	28.86
11	2.16	2.55	4.93	5.04	8.13	11.84	13.57	17.39	16.57	24.01
12	2.63	5.85	5.87	5.81	12.21	9.93	12.31	15.90	17.85	22.45

one break at most. Four subjects either did not exhibit a break in their data or the fit did not improve noticeably on going from two to three parameters. For these subjects the data were fitted without any break.

Table 3 gives the exponents, the ratios between scale units of the upper over the lower segment, and the positions of the breaks. The exponents range from .53 to 1.12. Figure 5 gives the plots.

Ratio-Setting Experiments

For all 12 subjects individually, and all three experiments, Φ_V was plotted against $(\Phi_S + \Phi_V)$, where Φ_S is the standard and Φ_V the variable duration (from the experimenter's point of view). This way of plotting the data agrees with Model C. Neither the short pause of 300 msec nor the reaction time of about 300 msec was taken into account. These intervals affect only the subjective zero(s), which are of minor interest

in this connection.⁵ In most cases the plots clearly showed whether one, two, or three straight lines should be fitted. The fitting was done with the method of least squares.

Denoting $\Phi_S + \Phi_V$ by Φ , we have, according to Model C,

$$r[\alpha(\Phi - \Phi_0)^\beta - (\Phi_V - \Phi'_0)^\beta] = (\Phi_V - \Phi'_0)^\beta, \quad (19)$$

$$\Phi_V = \left(\frac{r}{r+1} \right)^{1/\beta} \alpha^{1/\beta} \Phi - \left(\frac{r}{r+1} \right)^{1/\beta} \alpha^{1/\beta} \Phi_0 + \Phi'_0.$$

⁵ Allan and Kristofferson (1974) cite two duration discrimination studies where performance deteriorated with ISIs shorter than 500 msec. A comparison with the present experiments, with an ISI of 300 msec, is hardly justified, however, because both the tasks and the sets of durations used were different. Perhaps it can be concluded that Sensory Register 1 was not "switched off" during the short pause in the present experiments.

TABLE 3
EXPONENTS β AND RATIOS α OF SCALE UNIT OF UPPER SEGMENT OVER LOWER SEGMENT, WITH POSITION OF
BREAK IN SECONDS FOR FOUR EXPERIMENTS

Sub- ject	Magnitude estimation						Equal setting						Half setting						Double setting								
	β	α	Break	β	α	Break	No. L	Stress	No. P	β	α	Break	No. L	Stress	No. P	β	α	Break	No. L	Stress	No. P	β	α	Break	No. L	Stress	No. P
1	.579	1.718	3	.595	1.234	4.1	3	.070	10	.701	1.167	2.3	3	.047	10	.562	.560	10	2 ^a	.073	9						
2	.884	.844	4	.742	1.254	4.7	3	.037	8	.969	1.194	3.5	3	.052	10	.769	.724	16	2 ^b	.036	9						
3	.525	—	—	.619	1.116	(10)	2	.034	10	1.012	—	19	2	.098	10	.561	—	—	1	.062	10						
4	.920	1.370	2	.785	.738	12.5	2	.030	9	.768	1.143	12	2	.029	10	.549	.931	(12)	3	.028	10						
5	.729	2.456	13	.801	.729	(15)	2	.017	10	.721	—	—	1	.101	10	.701	.919	(21)	3 ^c	.054	9						
6	.834	2.227	2	.740	1.367	(3)	2	.024	10	.904	1.382	(4)	2	.044	10	.870	—	12.6	3 ^d	.053	8						
7	.948	.812	5	1.045	.577	7.8	3	.054	10	.816	—	—	1	.162	10	1.177	.555	(15)	2	.026	8						
8	1.009	—	—	1.009	—	—	1	.022	10	.989	—	—	1	.076	10	.947	.924	22	3	.028	10						
9	.746	—	—	.821	.449	12	2	.009	8	.845	.818	4.1	2 ^e	.035	8	.614	.782	11	3	.028	8						
10	.913	1.262	5	.905	.552	24	2	.027	10	1.294	.545	20	2	.014	10	.953	—	46	1 ^f	.045	9						
11	.943	—	—	.969	.548	13.2	2	.029	9	.823	—	23	1 ^f	.075	9	1.034	.671	(20)	2	.081	10						
12	1.121	.621	7	1.002	.602	10.8	2	.013	10	.852	1.795	2.7	3	.077	9	.818	.747	22	2	.118	10						

Note. Unsegmented psychophysical functions are indicated by— in the Break column. The position of the break is given with decimals only if the exact position could be determined. Values within parentheses denote overlapping segments. For the ratio-setting experiments, the number of lines (No. L) in the ϕ_r - ϕ_r plot, the stress, and the number of points on which the stress is based (No. P) are also given.

^a Plus 1 point in the upper break.

^b Plus 1 point in the uppermost line.

^c Transition line consists of only 1 point.

^d 2 points in the lower break, no transition line.

^e Plus 2 points in the upper break.

^f Plus 1 point in the transition line.

Setting $r/(r+1) = p$, for $r = .5, 1$, and 2 (HS, ES, and DS), we obtain $p = 1/3, 1/2$, and $2/3$, respectively.

For one straight line there were no problems. If there were two lines, they were fitted independently. The three-line case is somewhat more complicated. The two outer lines must be parallel, and the three intercepts are not independent (since there are only two segments, there are only two subjective zeros). Fitting all 10 points simultaneously with four parameters would have required the solution of a system of non-linear equations. This I did not consider worthwhile. Instead the two outer lines were fitted together (same slope), and the middle line separately. Let us consider the cases with one, two, and three straight lines one at a time.

For a single straight line we assume a single psychophysical function (no break) and obtain

$$\beta = \log p / \log a,$$

according to Equation 15, and Φ_0 from Equation 17.

For two straight lines the main problem is which of the two belongs to the same segment and which is the transition line between the two segments. Consider Figure 2. In order to have no overlap of the two segments of the psychophysical function (i.e., to have a unique mapping of Φ onto ψ), the highest ordinate of the straight line corresponding to the transition region (the middle line in Figure 2A) must not lie above the lowest abscissa of the same straight line (see Eisler, Note 2). Usually the data were unequivocal; in a few cases other criteria had to be invoked. From the line *not* belonging to the transition region, β and Φ_{01} were determined as described above; then α and Φ_{02} were computed from the other line:

$$\alpha = a'^{\beta/p}, \quad \Phi_{02} = b' + a'\Phi_{01},$$

or

$$\Phi_{02} = (\Phi_{01} - b')/a',$$

depending on whether the transition line lay below or above the other line. (Subscripts 1 and 2 for Φ_0 refer to the two segments;

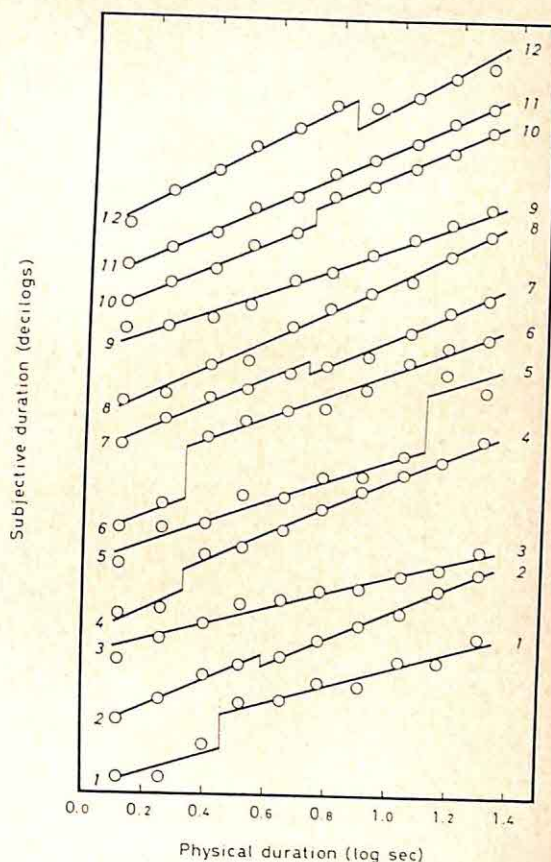


FIGURE 5. Log magnitude estimates as functions of log stimulus values for 12 subjects.

slope and intercept for the transition line are denoted by '.')

For three straight lines the exponent was determined as above from the slope of the outer lines, α , and α' as above from the slope of the middle line, α' . The intercepts of the outer lines yielded the subjective zeros.

Knowing the exponent β , the ratio between the scale units α , and the Φ_{0s} , the psychophysical function was computed. From the parameters of the straight line(s) the theoretical Φ values were computed. From the latter and the corresponding power function the ψ values were computed and, by dividing them by p , the ψ_r values. The ψ_r values in their turn were entered into the relevant power function (same or lower segment) and the corresponding Φ_r values were calculated. The points in Figures 2, 6, and 7 are these ψ (open circles) and ψ_r (filled circles) values plotted against the empirical Φ and Φ_r . Finally, the "stress," a normalized measure of

TABLE 5
 "EXPERIENCED" VALUES OF $\frac{1}{2}$ AND 2

Value	Subject											
	1	2	3	4	5	6	7	8	9	10	11	12
$\frac{1}{2}$	.65	.76	1.04	.48	.42	.69	.32	.48	.52	.87	.38	.38
2	1.87	2.09	1.77	1.27	1.70	2.43	2.31	1.85	1.39	2.13	2.16	1.56

Note. The exponent from the equal-setting experiment is assumed to be correct and the same as the exponents from the other two experiments.

ment between ES and DS is also good, but the correlation of β between ES and HS is low. The latter result is mainly attributable to two subjects, 3 and 10. One possible interpretation of the discrepancies between ES and HS and DS is a misconception by a subject about the numbers $1/2$ and 2. Table 5 gives the "experienced" numbers, assuming the exponent from ES to be the correct one. However, if a subject misconstrues a number in ratio setting, it is surprising that he does well in ME.

The most dependable experiments probably are ES and ME, giving average exponents of .84 and .85, respectively, though with an interindividual variation of .59–1.04 for ES and .53–1.12 for ME.

A by-product of these results is the conclusion that—at least for the continuum of time—different individual exponents reflect perceptual differences rather than differences in numerical habits, since ES is number

free (cf. Ekman, Hosman, Lindman, Ljungberg, & Åkesson, 1968; Jones & Marcus, 1961; Jones & Woskow, 1962; Rule, 1966, 1968, 1969; Rule & Markley, 1971; Wanschura & Dawson, 1974).

To test whether the ratio between scale units depends on subjects and tasks, an analysis of variance like the previous one was carried out with α as the dependent variable (for cases with only one segment, α was set at 1). The outcome was the reverse of the previous one: no significant differences for subjects, but significant differences for tasks ($p < .01$). Pairwise t testing showed that the tasks can be partitioned into two groups, namely, ES and DS on the one hand, and ME and HS on the other ($p < .05$ for all pairs except ES-DS and ME-HS). This result can be explained by a centering tendency: When the total physical duration is comparatively long, as in ES and DS, the subjective duration is compressed by a scale

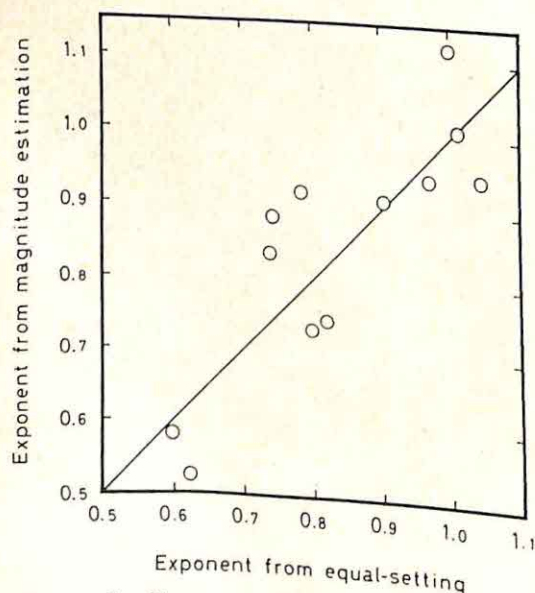


FIGURE 8. Exponents from magnitude estimation plotted against exponents from equal setting for 12 subjects.

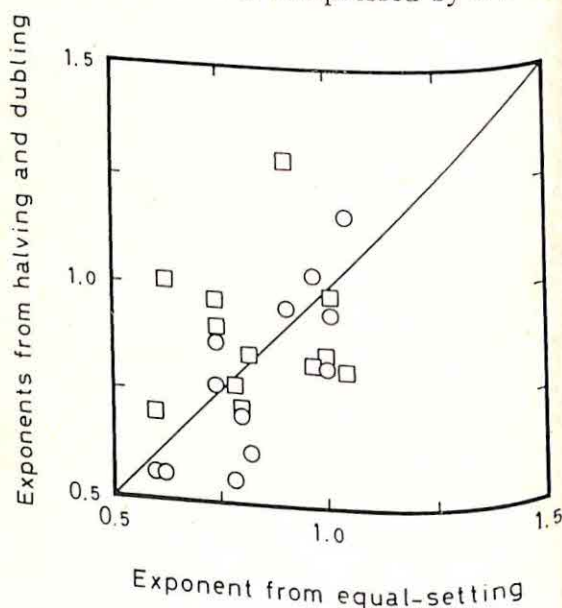


FIGURE 9. Exponents from half setting (squares) and double setting (circles) plotted against exponents from equal setting.

unit less than unity for the longest durations (the upper segment); but for comparatively short durations, as in ME and HS, it is expanded by a scale unit greater than unity for the longest durations.

EXPERIMENTAL RESULTS AND THE THREE MODELS

The three models cannot be distinguished on the basis of goodness of fit of the data. The fit ultimately depends solely on how well the experimental points fit the straight line or lines (see Figure 2A), so no statistical test can help in choosing.

The comparison of the three models is based on three considerations: (a) the slope of the Φ_r - Φ plots, (b) the breaks in these plots, and (c) the agreement of the exponents between the four scaling experiments.

The Slopes of the Φ_r - Φ Plots

This section is concerned only with Model A and the slopes for which standard and variable lie on the same segment of the psychophysical function. There Φ_r is seen as a function of Φ_s with the following requirements: The slope c for ES has to be unity; for DS it has to exceed unity, and for HS it must not exceed unity, since $\beta = \log r / \log c$ (Equation 13). For ES, $r = 1$; for DS, $r = 2$; and for HS, $r = 1/2$; entailing $c = 1$ to obtain a finite β for ES, and the fulfillment of the requirements mentioned for DS and HS to obtain $\beta > 0$.

Model A taken as it is. For ES the slope differs from unity for all subjects except three or four; for example, the slope for ES for Subject 1, when plotted in this way rather than as in Figure 2A, amounts to .45. Furthermore, for DS the slopes for Subjects 1, 3, and 4 are .95, .94, and .91, respectively. Thus the simple Model A has to be rejected.

Model A corrected for the regression effect. The regression effect described by Stevens and Greenbaum (1966) might allow a correction. If we assume that the "true" slopes of ES are unity, the slopes obtained may be attributed to this effect. This amounts to introducing a new parameter (at variance with the parsimony rule). Dividing the slopes for DS by the slopes for ES renders all DS slopes > 1 . However, if the same correction is carried out for the HS

data, the slopes for Subjects 2, 3, 6, and 10 exceed unity (1.21, 1.56, 1.08, and 1.61, respectively). Thus, this version of Model A also has to be rejected. Assuming different regression effects for the different tasks would, besides adding two more parameters, make the parameters unidentifiable, and a model with identifiable parameters (B or C in this case) is to be preferred.

Model A with different scale units for standard and variable. If we assume that the scale units of standard and variable always are different, that is, that standard and variable always lie on different segments, we would have to introduce one more parameter for each ratio-setting experiment; instead of the two segments in Figure 2B, we would have four. As in the corrected Model A above, the parameters were unidentifiable, since the exponent and the ratio of the scale units are linked (see Equation 13). And in the same manner as for the corrected Model A, this would be at variance both with the parsimony rule and the rule to choose models with identifiable parameters. Therefore, Model A is rejected (in all three versions).

The Breaks in the Φ_r - Φ Plots

Consider Figure 2. As mentioned before, the discontinuity in the ψ - Φ plot is situated at 4 sec, entailing a discontinuity at 4 sec for the x-axis and at the same value for the y-axis in the Φ_r - Φ plot. If the data were plotted according to Model A, Φ_r as a function of Φ_s , the break at the y-axis would still be at 4 sec, but the break at the x-axis at about 3 sec. Again, Model A is at variance with the data. We will not deal further with Model A since it has already been rejected; rather we will investigate whether the position of the breaks can be used in deciding between Models B and C.

As pointed out in the section on data analyses, with ratio settings, in order not to have an overlap in the psychophysical function (a unique mapping of ψ onto Φ), the highest ordinate of the straight line corresponding to the transition region (in Figure 2A it is the middle line) must not lie above the lowest abscissa of the same straight line.

For Model B the plot corresponding to Figure 2A has Φ_s as ordinate and the total

time (standard + variable) Φ_T as abscissa. For HS (which is most sensitive to the above condition) out of the eight subjects showing a break in the psychophysical function, seven violated the condition. (For Model C, the same condition was violated by only one subject.) Therefore, Model C is to be preferred to Model B.

Agreement of Exponents Between the Four Scaling Experiments

We have already seen that, when treating the data in accordance with Model C, we obtain high correlations (with one exception) over subjects between experiments. The correlation between ES and ME, the most interesting one, amounts to almost .9.

This gives further support for the rejection of Model A, since for the possible versions of this model the exponent is not identifiable. Model C is thus to be preferred to Model A not only because of the identifiability of the exponent as such, but also because its value agrees among experiments and thus carries real information.

When the exponent is computed in accordance with Model B, the correlation between ES and ME becomes $-.9$, likewise supporting the earlier decision to reject Model B.⁷ Therefore, Models A and B are rejected; Model C is retained.

THE TIME-ORDER ERROR

The explanation of the time-order error (TOE) is a simple consequence of the present model. Rewriting Equation 19 with $\Phi_V = \Phi_E$ (equal settings), $\Phi = \Phi_S + \Phi_E$ (Φ_S standard duration), and $r = 1$, we obtain

$$\alpha(\Phi_S + \Phi_E - \Phi_0)^\beta - (\Phi_E - \Phi'_0)^\beta = (\Phi_E - \Phi'_0)^\beta.$$

Rearranging yields

$$\text{TOE} = \Phi_E - \Phi_S = \frac{(2/\alpha)^{1/\beta} \Phi'_0 - \Phi_0 - [(2/\alpha)^{1/\beta} - 2] \Phi_S}{(2/\alpha)^{1/\beta} - 1}. \quad (20)$$

We see that the TOE depends, both in sign and size, on the interaction of the scale units and the subjective zeros on the one hand, and the standard duration on the other.

⁷ In a somewhat complicated and farfetched manner this correlation could be explained as a result of the "regression tendency" (Stevens, 1959; Stevens & Greenbaum, 1966).

Small wonder that it has been impossible to pin down and that different researchers obtain different results (Woodrow, 1951) or the same researchers different results for different subjects (Allan & Kristofferson, 1974; Allan, Kristofferson, & Rice, 1974). However, certain general features are apparent.

Consider the simple case with only one segment, that is, $\alpha = 1$, $\Phi'_0 = \Phi_0$. The denominator in Equation 20 is positive, and since typically $\beta < 1$, the third term of the numerator is negative (the expression between brackets positive). For large standards, Φ_0 can be neglected and the TOE is negative, in agreement with the common findings. Compare Φ and Φ_1 in Figure 3. For small standards the TOE becomes positive if $\Phi_S < [(2^{1/\beta} - 1)\Phi_0]/(2^{1/\beta} - 2)$. This agrees with the finding that if there is a positive TOE, it is for small standards—a positive TOE for small and a negative TOE for large durations is called Vierordt's law (Vierordt, 1868). With knowledge of the parameters of the psychophysical function, the point or points of indifference can be predicted exactly.

This argument is illustrated in Figure 10, showing data obtained in an equal-setting experiment by Fraisse (1948), where Φ_V is plotted against Φ_T . The thin line— $\Phi_V = \Phi_T/2$ —indicates veridical time perception

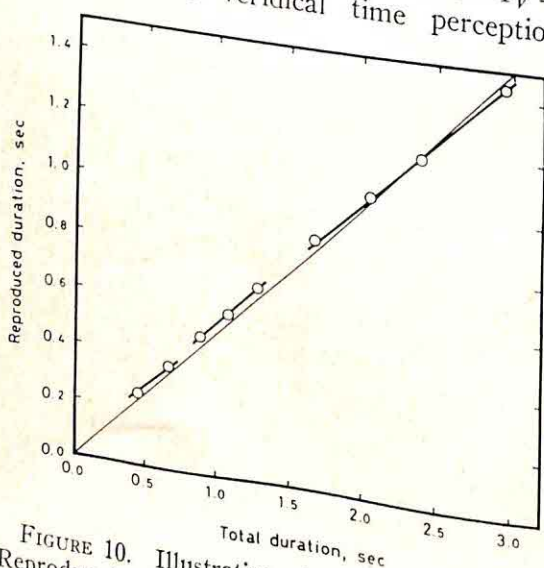


FIGURE 10. Illustration of the time-order error. Reproduced durations are plotted against the sum of standard and reproduced durations. The thin line represents veridical time perception. Data from Fraisse (1948).

with the variable duration equal to half the total duration. For durations exceeding 1.18 sec (the indifference point) the TOE is negative, since the slope of the straight line (the uppermost segment) is .4489 ($\beta = .87$), and thus smaller than .5. For shorter durations the TOE is positive because of the positive intercept of .12, corresponding to $\Phi_0 = .22$.

It is seen that the TOE for time is a special case. For other continua it is probably possible to describe the TOE in terms of different subjective zeros, and perhaps different scale units for standard and variable (cf. Eisler & Ottander, 1963). This may indicate that it might be more rewarding to study the variation of the parameters of the power law with experimental conditions than to study the TOE directly.

SUMMARY AND CONCLUSION

We have seen that Model C copes with the problem of time continuing to elapse after the presentation of the first duration of a pair, as well as with the problem of insufficient time for coding and transferring to LTM the first duration for short ISIs (by eliminating any memory), and that Model C is capable of accounting for features in discrimination characteristic for durations only. Furthermore, only Model C can account satisfactorily for the data obtained in the four scaling experiments. Note that the same process seems to be at work in discrimination studies, where usually short durations are used, and in scaling experiments with comparatively long durations.

An instance of particular interest is one of the experiments carried out by Richards (1964), because it covers a range from 2 sec to 80 min (standards; the smallest response time is 2.57 sec and the largest total time 10,400 sec) and involved one subject. Here we find two breaks; one at 10 sec and the other at 600 sec. Accordingly, points 4-7 and 10-12 of the 12 data points can be used to determine the exponent; the two slopes are .574 and .559, respectively, a difference of only .015, corresponding to a difference of the exponents of .06.

Data on duration from other investigators can be found in Eisler (Note 3). Again,

Model C is the only one of the three models that yields satisfactory results for ratio-setting experiments throughout. The exponents obtained agree for the majority of these studies with the exponents found in the present series of experiments, thus lying somewhat below the value of unity given by Stevens (e.g., 1967).

As a result of the data analysis in combination with Model C, it proved possible to reject psychophysical functions other than Stevens' power law, and to obtain exponents for duration from equal setting, that is without having to rely on a subject's numerical behavior. The latter finding not only obviates objections to direct scales as being artifacts of the subject's numerical habits; it also opens up opportunities to obtain exponents for duration from animal experiments in a more convincing way than, for example, in the experiment of Platt, Kuch, and Bitgood (1973). An experiment of equal setting of duration with rats has in fact been carried out with the same equipment and the same stimuli as described above; it will be reported in due course.

The possibility of determining exponents from ES data also allows comparisons between different modalities, used for presenting the time intervals, in a more precise manner. The same goes for comparisons between different groups of subjects, though remembering that the exponents for the presumably normal subjects in the present ES experiment varied between .6 and 1.0, the group differences found by Carlson and Feinberg (1968), for example, between normals ($\beta = .96$) and schizophrenics ($\beta = .91$) seem rather unimpressive. In this connection the importance of treating the data individually should be emphasized.

Model C has also made it possible to describe the TOE for durations in a parametric way; a task for future research may be to investigate in detail which factors affect the location of the break in the psychophysical function for duration and the values of the three relevant parameters: the exponent, the scale unit, and the subjective zero.

REFERENCE NOTES

1. Ekman, G., & Frankenhaeuser, M. Subjective time scales (Report No. 49). Stockholm: Re

experiments, the reader should consult Ikeda and Obonai (1955) and Martin (1974a).³

In the theoretical interpretation of experiments of this type, it is generally assumed that there are neurons that respond to the target when it is presented alone and neurons that respond to the inducer when it is presented alone. (The two sets of neurons are not necessarily disjoint.) The task for theory, then, is to explain why the occurrence of responses in the inducer neurons results in a modification of responses in neurons associated with the percept of the target or in some modification of the overall neural responding such that the target signal cannot be as easily extracted from the noise. This way of thinking certainly seems plausible. Indeed, at first sight, it seems so compelling that it might appear to exhaust the logical possibilities. Nonetheless, it is my purpose in this article to suggest that reasoning in this manner is unnecessarily restrictive. In effect, I want to raise the possibility that at least in some cases not two, but three classes of neurons could be responding to the two stimuli presented.

The three classes to which I am referring include the two that are ordinarily considered (the T or target neurons and the I or inducer neurons) and a third class that responds minimally, if at all, to either target or inducer alone. Neurons of this type respond, however, to their combined presentation at appropriate temporal intervals. Accordingly, they will be called T-I (target-inducer) neurons for the general case. When the inducer is a mask, the I neurons will be called M (mask) neurons, and the T-I neurons will be called T-M (target-mask) neurons. Anticipating the argument to be developed more carefully below, I simply point out for the moment that succession (motion) detectors are among the neurons that might be found in this third class and that in gen-

eral both the members of the T-I class stimulated on a given trial and the magnitude of response from the stimulated neurons depend on the temporal interval between target and inducer on that trial. To forestall any possible misunderstanding, however, it should be noted from the outset that calling these neurons motion detectors does not necessarily imply their involvement in processes leading to the sensation of movement. It is also important to note that the reference to three classes of neurons does not imply homogeneity within these classes;⁴ the three classes are not single channels in the sense in which that expression has been used in the recent literature (Weinstein, Ozog, & Szoc, 1975; Graham, in press).

Although the general thesis of this paper could apply to many two-transient experiments,⁵ I will illustrate the possibilities for thinking with the additional degree of freedom afforded by the target-inducer neurons with one particular case only. That case is lateral masking, in which target and mask fall on neighboring, but not overlapping, retinal areas. To facilitate the discussion, it is presented in two parts. The first of these deals simply with the plausibility of assuming that T-M neurons exist and are stimulated in lateral masking experiments, independently of any possible relevance they might have in theoretical interpretations. The second part of the discussion deals with some implications of the further assumption that masking results from the joint action of the M neurons and the T-M neurons on the T neurons.

⁴ Actually, several different types of neurons would be expected within each of the three classes. On-center cells, off-center cells, X cells, and Y cells that fire to the target or to the inducer are examples.

³ The two-transient paradigm is also employed in studies of apparent movement. However, the latter studies differ from the other experiments being considered in that they are not generally concerned with the measurement of changes in the sensation associated with one stimulus as the result of the presentation of another. Rather, their purpose is to study an entirely new sensation.

⁵ Whether the combined presentation of the stimuli used as target and inducer in any particular experiment leads to a response in T-I neurons is essentially an empirical question. I venture to guess, however, that minimal effects, if any, would be found when the target is a small patch of light concentric with a very much larger mask, as in the experiments of Sperling (1965) and of Boynton and his colleagues (Boynton, 1972). On the other hand, large effects might be expected in some types of lateral masking.

EXISTENCE OF TARGET-MASK NEURONS

Potentially, several types of neurons could fall into the T-M class in a lateral masking experiment. Cells with large receptive fields, for example, might not be optimally stimulated by either target or mask alone, but might respond favorably to their combined presentation. However, the most convincing evidence for the existence of T-M neurons is found in the literature on motion detectors. There it has been shown, for a considerable variety of animals, that successive stimulation of neighboring retinal regions, at appropriate temporal and spatial separations, is an adequate stimulus for a motion detector. These animals include several insects (McCann, 1973; Reichardt, 1961, 1969), rabbits (Barlow & Levick, 1965), and frogs (Grüsser-Cornehls, 1968; Grüsser-Cornehls, Grüsser, & Bullock, 1963). To my knowledge, no attempt has been made to study this matter in primates. However, it seems plausible to suggest that the motion detectors that have been found in that order of animals (Hubel & Wiesel, 1968; Wurtz, 1969a, 1969c) could be shown to be succession detectors with appropriate experiments. Then, if the human does not differ from other primates in this regard, it could be assumed that neurons exist in the human visual system, possibly at a variety of neural levels, which are specialized for the detection of relative motion between retina and environment but which also respond to the target-mask stimulus sequence in a lateral masking experiment.

The reader should note carefully, however, that relative motion between retina and environment does not necessarily imply the movement of an object; such motion could equally well arise from movements by the observer and, in particular, from the movements of his eyes. Indeed, beginning with an important series of studies in the striate cortex of the awake, unanesthetized monkey (Wurtz, 1968, 1969a, 1969b), a rapidly accumulating body of evidence has become available which shows that information about the motion of the retinal image during saccadic eye movements, including its high velocity components, is processed by

the visual system (Adey & Noda, 1973; Brooks & Holden, 1973; Hoffmann, 1973; Noda, 1975; Noda & Adey, 1974; Noda, Freeman, & Creutzfeldt, 1972; Noda, Freeman, Gies, & Creutzfeldt, 1971; Wurtz & Goldberg, 1972).

In addition to the studies cited in the preceding paragraph, which come specifically from the literature on saccadic eye movements, there are scattered references in the general neurophysiological literature to the existence at various neural levels of cells that respond to velocities greater than 100° of visual angle per sec. Such references can be found in studies specifically concerned with neurons responding to movement (e.g., Oyster, Takahashi, & Collewijn, 1972, Fig. 6). They can also be found in the literature on W, X and Y cells,⁶ in which it has been shown that some of the Y cells, but no X cells, respond to very high velocities of image motion (e.g., Cleland, Dubin, & Levick, 1971, p. 481; Cleland, Levick, & Sanderson, 1973, p. 652; Hughes & Whitteridge, 1973, p. 1105; Singer & Bedworth, 1973, pp. 296-297). It seems unlikely that cells responding to very high velocities are involved in the sensation of motion, because, at least in man, such sensations are only elicited by slow movements. However, a large range of velocities, including some very high ones, are quite routinely generated on the retina during saccadic eye movements. Accordingly, it seems reasonable to suggest that those broadly tuned Y cells which respond to high velocities of image motion are a part of the neural apparatus that is processing information related to saccadic eye movements or other rapid bodily movements, with the more peripheral cells of this type feeding into central cells sensitive to high velocities. As will be seen shortly, an important role will be assigned to these neurons in

⁶ The X-Y cell dichotomy was introduced by Enroth-Cugell and Robson (1966). These two cell types are sometimes also referred to as sustained-transient (Cleland, Dubin, & Levick, 1971) and Type II, Type I (Fukada, 1971). More recently, a third category (W) has been introduced (Stone & Hoffmann, 1972).

experiments, the reader should consult Ikeda and Obonai (1955) and Matin (1974a).³

In the theoretical interpretation of experiments of this type, it is generally assumed that there are neurons that respond to the target when it is presented alone and neurons that respond to the inducer when it is presented alone. (The two sets of neurons are not necessarily disjoint.) The task for theory, then, is to explain why the occurrence of responses in the inducer neurons results in a modification of responses in neurons associated with the percept of the target or in some modification of the overall neural responding such that the target signal cannot be as easily extracted from the noise. This way of thinking certainly seems plausible. Indeed, at first sight, it seems so compelling that it might appear to exhaust the logical possibilities. Nonetheless, it is my purpose in this article to suggest that reasoning in this manner is unnecessarily restrictive. In effect, I want to raise the possibility that at least in some cases not two, but three classes of neurons could be responding to the two stimuli presented.

The three classes to which I am referring include the two that are ordinarily considered (the T or target neurons and the I or inducer neurons) and a third class that responds minimally, if at all, to either target or inducer alone. Neurons of this type respond, however, to their combined presentation at appropriate temporal intervals. Accordingly, they will be called T-I (target-inducer) neurons for the general case. When the inducer is a mask, the I neurons will be called M (mask) neurons, and the T-I neurons will be called T-M (target-mask) neurons. Anticipating the argument to be developed more carefully below, I simply point out for the moment that succession (motion) detectors are among the neurons that might be found in this third class and that in gen-

eral both the members of the T-I class stimulated on a given trial and the magnitude of response from the stimulated neurons depend on the temporal interval between target and inducer on that trial. To forestall any possible misunderstanding, however, it should be noted from the outset that calling these neurons motion detectors does not necessarily imply their involvement in processes leading to the sensation of movement. It is also important to note that the reference to three classes of neurons does not imply homogeneity within these classes;⁴ the three classes are not single channels in the sense in which that expression has been used in the recent literature (Weisstein, Ozog, & Szoc, 1975; Graham, in press).

Although the general thesis of this paper could apply to many two-transient experiments,⁵ I will illustrate the possibilities for thinking with the additional degree of freedom afforded by the target-inducer neurons with one particular case only. That case is lateral masking, in which target and mask fall on neighboring, but not overlapping, retinal areas. To facilitate the discussion, it is presented in two parts. The first of these deals simply with the plausibility of assuming that T-M neurons exist and are stimulated in lateral masking experiments, independently of any possible relevance they might have in theoretical interpretations. The second part of the discussion deals with some implications of the further assumption that masking results from the joint action of the M neurons and the T-M neurons on the T neurons.

⁴ Actually, several different types of neurons would be expected within each of the three classes. On-center cells, off-center cells, X cells, and Y cells that fire to the target or to the inducer are examples.

⁵ Whether the combined presentation of the stimuli used as target and inducer in any particular experiment leads to a response in T-I neurons is essentially an empirical question. I venture to guess, however, that minimal effects, if any, would be found when the target is a small patch of light concentric with a very much larger mask, as in the experiments of Sperling (1965) and of Boynton and his colleagues (Boynton, 1972). On the other hand, large effects might be expected in some types of lateral masking.

³ The two-transient paradigm is also employed in studies of apparent movement. However, the latter studies differ from the other experiments being considered in that they are not generally concerned with the measurement of changes in the result of the presentation of another. Rather, their purpose is to study an entirely new sensation.

EXISTENCE OF TARGET-MASK NEURONS

Potentially, several types of neurons could fall into the T-M class in a lateral masking experiment. Cells with large receptive fields, for example, might not be optimally stimulated by either target or mask alone, but might respond favorably to their combined presentation. However, the most convincing evidence for the existence of T-M neurons is found in the literature on motion detectors. There it has been shown, for a considerable variety of animals, that successive stimulation of neighboring retinal regions, at appropriate temporal and spatial separations, is an adequate stimulus for a motion detector. These animals include several insects (McCann, 1973; Reichardt, 1961, 1969), rabbits (Barlow & Levick, 1965), and frogs (Grüsser-Cornehls, 1968; Grüsser-Cornehls, Grüsser, & Bullock, 1963). To my knowledge, no attempt has been made to study this matter in primates. However, it seems plausible to suggest that the motion detectors that have been found in that order of animals (Hubel & Wiesel, 1968; Wurtz, 1969a, 1969c) could be shown to be succession detectors with appropriate experiments. Then, if the human does not differ from other primates in this regard, it could be assumed that neurons exist in the human visual system, possibly at a variety of neural levels, which are specialized for the detection of relative motion between retina and environment but which also respond to the target-mask stimulus sequence in a lateral masking experiment.

The reader should note carefully, however, that relative motion between retina and environment does not necessarily imply the movement of an object; such motion could equally well arise from movements by the observer and, in particular, from the movements of his eyes. Indeed, beginning with an important series of studies in the striate cortex of the awake, unanesthetized monkey (Wurtz, 1968, 1969a, 1969b), a rapidly accumulating body of evidence has become available which shows that information about the motion of the retinal image during saccadic eye movements, including its high velocity components, is processed by

the visual system (Adey & Noda, 1973; Brooks & Holden, 1973; Hoffmann, 1973; Noda, 1975; Noda & Adey, 1974; Noda, Freeman, & Creutzfeldt, 1972; Noda, Freeman, Gies, & Creutzfeldt, 1971; Wurtz & Goldberg, 1972).

In addition to the studies cited in the preceding paragraph, which come specifically from the literature on saccadic eye movements, there are scattered references in the general neurophysiological literature to the existence at various neural levels of cells that respond to velocities greater than 100° of visual angle per sec. Such references can be found in studies specifically concerned with neurons responding to movement (e.g., Oyster, Takahashi, & Collewijn, 1972, Fig. 6). They can also be found in the literature on W, X and Y cells,⁶ in which it has been shown that some of the Y cells, but no X cells, respond to very high velocities of image motion (e.g., Cleland, Dubin, & Levick, 1971, p. 481; Cleland, Levick, & Sanderson, 1973, p. 652; Hughes & Whitteridge, 1973, p. 1105; Singer & Bedworth, 1973, pp. 296-297). It seems unlikely that cells responding to very high velocities are involved in the sensation of motion, because, at least in man, such sensations are only elicited by slow movements. However, a large range of velocities, including some very high ones, are quite routinely generated on the retina during saccadic eye movements. Accordingly, it seems reasonable to suggest that those broadly tuned Y cells which respond to high velocities of image motion are a part of the neural apparatus that is processing information related to saccadic eye movements or other rapid bodily movements, with the more peripheral cells of this type feeding into central cells sensitive to high velocities. As will be seen shortly, an important role will be assigned to these neurons in

⁶ The X-Y cell dichotomy was introduced by Enroth-Cugell and Robson (1966). These two cell types are sometimes also referred to as sustained-transient (Cleland, Dubin, & Levick, 1971) and Type II, Type I (Fukada, 1971). More recently, a third category (W) has been introduced (Stone & Hoffmann, 1972).

the suggestions about lateral masking that are outlined below.⁷

MASKING AS A JOINT RESULT OF MASK AND TARGET-MASK NEURONS

If the assumption that T-M neurons are stimulated in a lateral masking experiment is correct, we are confronted with three logically possible alternatives: (a) The responses of the T-M neurons could be irrelevant to the masking effect, which would be due entirely to the M neurons; (b) all T-M neurons could contribute to masking; or (c) some, but not all T-M neurons could contribute to it. None of these alternatives can be accepted or rejected out of hand. However, I choose the third for further exploration and, in effect, follow earlier suggestions that some motion detectors could be involved in the sensation of movement and others in saccadic suppression and visual masking (Matin, 1974b, in press). The discussion begins with a brief summary of that information about lateral masking which is necessary for the subsequent analysis.

Some Pertinent Facts

When lateral masking is studied, the backward temporal intervals (target precedes mask) are called metacontrast and the forward intervals paracontrast. In general, and a priori perhaps somewhat surprisingly, there is a gross asymmetry in magnitude of effect in these two cases. Whereas paracontrast is usually a relatively weak phenomenon, metacontrast can be very powerful: Under appropriate experimental conditions, stimuli that are clearly visible when presented alone can undergo marked decreases in brightness or can even be suppressed entirely (Alpern, 1953; Fry, 1934;

Schiller & Smith, 1966; Werner, 1935). Frequently, along with the changes in brightness, subjects report an apparent movement of target toward the mask (Fehrer, 1965, 1966; Kahneman, 1967; Kolers & Rosner, 1960; Toch, 1956).

Under some conditions, the magnitude of metacontrast is a monotonic decreasing function of time between the onset of the target and the onset of the mask. Under others, masking is maximal when the target precedes the mask by some interval; the obtained temporal functions are then called U-shaped. Depending again on the particular conditions employed, this interval can vary from just a few msec to more than 100 msec. When the observer's task is the recognition or detection of the target, there is considerable controversy about the shape of the masking function: Some experimenters report monotonic functions⁸ and others monotonic functions or U-shaped functions, depending on the experimental conditions. However, there is no disagreement about the fact that metacontrast functions are U-shaped when a target that is clearly supra-threshold and of moderate energy is followed by a mask of similar energy, with the subject instructed to report on the brightness of the target⁹ (See Kahneman, 1968, and Lefton, 1972, for references). In many ways, this latter case of metacontrast is the most puzzling from a theoretical point of view. A reader who is unfamiliar with the masking literature might be inclined to overlook the proviso that mask and target must be of similar energy and therefore could be misled into thinking that a simple explanation for U-shaped functions could be found in the latency-intensity effect. That is, it might seem that a high-energy mask will

⁸ Those who are not familiar with the masking literature are advised that the expression *monotonic function* refers only to forward masking or backward masking alone. No masking function is monotonic when the entire possible range of temporal intervals is studied.

⁹ For the purpose of this discussion, no distinction is made between masking that involves decrements in brightness and masking as it is studied in detection experiments. However, the reader is advised that such a distinction should be considered. Consult Bridgeman (1975, p. 96), Erikson, Becker, and Hoffman (1970, p. 245), and Weinstein (1972, p. 235).

⁷ Implicit in this discussion is the assumption that all motion-sensitive cells, including the Y cells, are succession detectors. Although this assumption is plausible, it goes beyond the established facts and it would certainly be desirable to study more kinds of motion-sensitive cells with stimuli delivered successively to neighboring retinal areas. Note also that in this paper the terms *motion detectors* and *motion-sensitive cells* are used very loosely to indicate cells that are sensitive to movement. These neurons might include the more stringently defined direction-specific cells considered by some authors (e.g., Barlow & Hill, 1963), but they might also include other types of cells.

produce a neural response at a shorter latency than a target of lower energy and that this difference in latency could account for the apparently retroactive effect of mask on target. I therefore emphasize that the facts do not support such an interpretation: If mask and target energies are held fixed throughout the range of temporal intervals explored, the temporal functions are U-shaped when mask and target are of similar energy and monotonic when the mask is considerably greater in energy than the target (Kolars, 1962; Weisstein, 1968, 1972). On the other hand, a variety of other explanations have been proposed for U-shaped functions (Breitmeyer, Love, & Wepman, 1974; Bridgeman, 1971, 1975; Fehrer, 1965, 1966; Kahneman, 1967; Mayzner, Tresselt, & Helfer, 1967; Weisstein, 1968, 1972; Weisstein et al., 1975; Werner, 1935). As will be seen in subsequent paragraphs, the present article will produce an addition to this already long list of suggestions.

Apparent Movement and Metaccontrast

As was noted previously, several experimenters have commented on a correlation between apparent movement and masking in metaccontrast experiments (Biederman-Thorson, Thorson, & Lange, 1971; Breitmeyer, Love, & Wepman, 1974; Fehrer, 1965, 1966; Fehrer & Raab, 1962; Kahneman, 1967; Kolars & Rosner, 1960; Toch, 1956). In addition, a theory has been developed based on the assumption that this correlation is causal (Kahneman, 1967). According to this theory, metaccontrast is a case of impossible apparent movement. Thus, in the frequently employed display with three rectangles (a central one as target and two flanking ones as a mask), a single object is set into apparent motion in two directions—a physically impossible movement. Similarly, in the disc-ring display, the stimulus sequence presented would cause the original disc to grow larger even as it disappears. Kahneman suggested that suppression of the target occurs because the perceptual system is unable to resolve these contradictions. He supported his contention of a causal relation between masking and apparent movement with an experiment in which these two phenomena were studied

with the same subjects and similar displays. Under his conditions, the temporal functions for apparent movement and metaccontrast were virtually identical.

Although there are many difficulties with Kahneman's theory, it might have been possible to reformulate it in a satisfactory way if further evidence had substantiated the high correlation between temporal functions for apparent movement and masking that it implies. This, however, was not the case. It was shown, for example, that the resemblance between them breaks down when visual angle, luminance, and duration are explored over a sufficient range (Weisstein & Growney, 1969). Kolars also emphasized that the two phenomena are different in basic ways. Among other differences, he noted that masking falls off sharply as distance between target and mask is increased, but apparent movement does not, and that masking is markedly sensitive to similarity of features in target and mask, but apparent movement is not (Kolars, 1972, pp. 131-132).

Clearly, then, there is no causal relation between apparent movement and masking. Nonetheless, the intriguing correlation between the two effects remains and needs to be accounted for. The T-M neurons, I think, offer considerable promise toward that end. Recall the previously cited evidence which shows that at least some movement detectors respond to velocities well above those eliciting the sensation of motion in man. As was noted, such velocities would normally be produced in the course of saccadic eye movements. Let us suppose, then, that a variety of broadly tuned neurons exist, which detect relative motion between retina and environment. (Such neurons are not, of course, first-order neurons, nor need they be retinal neurons at all. Moreover, there is no reason to assume that they exist at only one neural level.)¹⁰ Some of

¹⁰ The cells referred to herein as *target neurons* and those referred to as *mask neurons* include first-order neurons and various other retinal and central neurons, presumably organized in some hierarchal manner as the visual message proceeds from the periphery through the central nervous system. On the other hand, because the types of T-M cells being considered in this paper consist of motion detectors, they do not include first-order

masking effects by the M neurons occur at the peripheral retina (e.g., Boynton, 1972), they could preclude the initiation of responding in the T-M neurons (which are not first-order neurons) and thus produce masking functions due entirely to M neurons.

Considerations about the T-M neurons might also be helpful in resolving the controversy in the masking literature about the existence of U-shaped functions when the observer's task is detection or recognition of the target (see Kahneman, 1968, and Lefton, 1972, for references). In experiments of this type, stimuli employed as targets tend to be weak. Therefore, differences in experimental conditions that might appear slight and apparently insignificant could result in a firing of T-M neurons in some studies but not in others, and therefore to reports of U-shaped functions in some studies but not in others (cf. Kahneman, 1968, p. 419).

Paracontrast

Because responses in the T-M neurons are elicited by successive stimulation of neighboring retinal areas, it could be expected that firing would occur in these neurons to the paracontrast sequence as well as the metacontrast sequence. At first sight, therefore, it might seem that the suggestions being considered would lead to a prediction of masking functions that are symmetrical around the origin (simultaneous presentation of target and mask). Such a prediction, of course, would be contrary to fact. In those cases where metacontrast is U-shaped, paracontrast is much weaker and does not extend as far from the origin (Alpern, 1953; Kolers & Rosner, 1960).

A more careful look at the situation, however, shows that masking as a joint result of the M and T-M neurons implies stimulus asymmetries between the forward and backward temporal intervals that do not exist for the mask stimulus alone, that is, stimulus asymmetries other than the mere fact that the target is flashed before the mask in metacontrast and the mask before the target in paracontrast. A fundamental temporal and

that are called U-shaped), latency-intensity considerations suggest that experiments of sufficient sensitivity would show peaks that are not exactly at zero.

a secondary spatial asymmetry for the T-M neurons are described in the next two paragraphs. This description is followed by a further discussion showing that a plausible explanation for the weakness of paracontrast relative to metacontrast can be derived from the fundamental asymmetry. In addition, I will raise the possibility that the secondary asymmetry might also contribute to a weaker paracontrast. The latter possibility, however, is by no means crucial to the overall argument.

Consider again what happens in a metacontrast experiment. The target is presented, followed by the mask. Accordingly, the target-mask stimulus sequence is completed when the mask is presented. For paracontrast, however, the completion of the target-mask stimulus sequence coincides with the presentation of the target, not the mask. In a sense, therefore, it might be said that the target-mask stimulus follows the target in metacontrast but is simultaneous with the target in paracontrast.

The asymmetry between metacontrast and paracontrast for the T-M neurons described in the preceding paragraph is inherent to the masking paradigm per se and independent of the particular stimuli that are chosen as target and mask. In addition, however, another secondary stimulus asymmetry exists in masking experiments as they are ordinarily performed. Consider the commonly used display in which the mask consists of two rectangles flanking a third rectangle, which serves as the target. For metacontrast, the stimulus for the T-M neurons consists of the presentation of a single, central rectangle followed by two flanking rectangles. For paracontrast, the situation is reversed. Thus, the stimuli for the T-M neurons are actually not the same in the forward and backward sequences. Essentially similar considerations apply to the other frequently used display, in which the target and mask are a disc and a ring, respectively.

In evaluating the possible effects of these two stimulus differences in producing the weakness of paracontrast relative to metacontrast, let us consider the secondary (spatial) asymmetry first. When the target is a rectangle and the mask two flanking rectangles, the metacontrast T-M sequence

consists of the single central rectangle, followed by two flanking ones; for paracontrast, the converse is true. I know of no convincing a priori reason to expect the paracontrast sequence to be less favorable for the stimulation of T-M neurons than the metacontrast sequence. Nonetheless, this possibility exists and should be explored. In any event, it seems extremely unlikely that the effect of the two different sequences could be identical, and it would clearly be desirable to perform quantitative studies with a simpler type of stimulus configuration without this asymmetry,¹⁶ that is, studies with two identical stimuli, one the mask, the other the target, located in such a manner that they fall on retinal tissues of similar sensitivity. In particular, a study with target and mask consisting of two parallel slits would be useful. In addition to eliminating the spatial differences between metacontrast and paracontrast for the T-M neurons, the use of such slits would have an additional advantage: It might tend to generate data that would facilitate an analysis of the suggested relation between masking and motion detectors because moving slits are frequently employed in the study of the latter.

Although the secondary stimulus asymmetry is removable by an appropriate choice of stimuli, the other, more fundamental asymmetry is not. Differences between metacontrast and paracontrast would therefore be expected even in experiments with simpler spatial configurations. Moreover, the considerations about the rapid conduction velocity of the T-M neurons relative to the T neurons that were raised in the previous discussion of metacontrast lead to the expectation that these differences would be in the direction of a weaker paracontrast. For the latter, the stimulus for the T-M neurons is completed when the target is presented. Therefore, if the T-M neurons conduct more rapidly than the T neurons, their effect might be at least partially dissipated by the time the responses of the T neurons arrive at the site of interaction. (Compare

this with the situation in metacontrast, for which the more rapid conduction velocity of the T-M neurons works in the direction of an increase in masking, because it gives the responses of the T-M neurons an opportunity to overtake the responses in the T neurons which were stimulated earlier.) Indeed, if the difference in conduction velocity between the T and the T-M neurons were great enough, the T neurons in a paracontrast experiment could escape the masking effect of the T-M neurons entirely. The paracontrast effect would then be the result of the M neurons alone. In either case, whether the escape is partial or complete, *a paracontrast that is weaker than metacontrast would be predicted.*

CONCLUDING REMARKS

As was shown in the preceding discussion, a considerable amount of information has become available in recent years about neurons in the visual nervous system that are responsive to rapid displacements of the retinal image. This information forms the empirical basis for the speculations about lateral masking in this article. In essence, I have raised the possibility that broadly tuned motion detectors, capable of responding to high velocities of image motion, are being tapped by the target-mask stimulus sequence in a lateral masking experiment, and I have explored a few of the implications of that assumption. The suggestions presented are not, and were not meant to be, a complete model of lateral masking. For example, the argument developed explicitly attributes some of the overall masking effect to neurons stimulated by the mask (i.e., to those neurons that are entirely responsible for the masking effect in other theoretical treatments). However, almost no attempt to examine the possible effects of these mask neurons was made.¹⁷ Nor does the paper even include a complete exploration of the

¹⁶ Such stimuli have occasionally been used in the masking literature (e.g., Stigler, 1910, who described masking experiments in which the target was half of a disc and the mask the other half). To my knowledge, however, no quantitative data are available on this question.

¹⁷ Nothing that was said in this article per se precludes any of the various masking mechanisms that have been previously proposed. For example, masking by the target-mask neurons could exist in addition to the masking of the X cells stimulated by the target by the Y cells stimulated by the mask, as has been suggested in recent theoretical treatments of metacontrast (Breitmeyer & Ganz, in press; Breitmeyer, Love, & Wepman, 1974; Weisstein et al., 1975).

effects that might be expected from the postulated target-mask neurons. Rather, my interest has been to use lateral masking to illustrate the more general thesis of this article—that three classes of neurons could sometimes be responding to the two stimuli presented in experiments employing the two-transient paradigm, and that theorists are unnecessarily limiting the choices available to them by overlooking this possibility. This general thesis need not be limited to experiments in visual lateral masking. Indeed, it need not be limited to masking experiments (or for that matter, to vision experiments) at all.

REFERENCE NOTE

1. Matin, E. *The two-transient (masking) paradigm*. Paper based on the contents of the present article and presented at the meeting of the Association for Research in Vision and Ophthalmology, Sarasota, Fla., May 1, 1975.

REFERENCES

- Adey, W., & Noda, H. Influence of eye movements on geniculostriate excitability in the cat. *Journal of Physiology*, 1973, 235, 805-821.
- Alpern, M. Metacontrast. *Journal of the Optical Society of America*, 1953, 43, 648-657.
- Barlow, H., & Hill, R. Selective sensitivity to direction of movement in ganglion cells of the rabbit retina. *Science*, 1963, 139, 412-414.
- Barlow, H., & Levick, W. The mechanism of directionally selective units in rabbit's retina. *Journal of Physiology*, 1965, 178, 477-504.
- Biederman-Thorson, M., Thorson, J., & Lange, G. Apparent movement due to closely spaced sequentially flashed dots in the human peripheral field of vision. *Vision Research*, 1971, 11, 889-903.
- Blanc-Garin, J. Le masquage visuel retroactif. *L'Année Psychologique*, 1967, 67, 153-193.
- Boynton, R. Discrimination of homogeneous double pulses of light. In D. Jameson & L. Hurvich (Eds.), *Handbook of sensory physiology* (Vol. 7, part 4: *Visual psychophysics*). Berlin: Springer-Verlag, 1972.
- Breitmeyer, B., & Ganz, L. Implications of sustained and transient channels for theories of visual pattern masking, saccadic suppression, and information processing. *Psychological Review*, in press.
- Breitmeyer, B., Love, R., & Wepman, B. Contour suppression during stroboscopic motion and metacontrast. *Vision Research*, 1974, 14, 1451-1456.
- Bridgeman, B. Metacontrast and lateral inhibition. *Psychological Review*, 1971, 78, 528-539.
- Bridgeman, B. Correlates of metacontrast in single cells of the cat visual system. *Vision Research*, 1975, 15, 91-99.
- Brooks, B., & Holden, A. Suppression of visual signals by rapid image displacement in the pigeon retina: A possible mechanism for saccadic suppression. *Vision Research*, 1973, 13, 1387-1390.
- Cleland, B., Dubin, M., & Levick, W. Sustained and transient neurones in the cat's retina and lateral geniculate nucleus. *Journal of Physiology*, 1971, 217, 473-496.
- Cleland, B., Levick, W., & Sanderson, K. Properties of sustained and transient ganglion cells in the cat retina. *Journal of Physiology*, 1973, 228, 649-680.
- Enroth-Cugell, C., & Robson, J. The contrast sensitivity of retinal ganglion cells of the cat. *Journal of Physiology*, 1966, 187, 517-552.
- Erickson, C., Becker, B., & Hoffman, J. Safari to masking land: A hunt for the elusive U. *Perception & Psychophysics*, 1970, 8, 245-250.
- Fehrer, E. Contribution of perceptual segregation to the relationship between stimulus similarity and backward masking. *Perceptual and Motor Skills*, 1965, 21, 27-33.
- Fehrer, E. Effect of stimulus similarity on retroactive masking. *Journal of Experimental Psychology*, 1966, 71, 612-615.
- Fehrer, E., & Raab, D. Reaction time to stimuli masked by metacontrast. *Journal of Experimental Psychology*, 1962, 63, 143-147.
- Frisby, J. The effect of stimulus orientation on the phi phenomenon. *Vision Research*, 1972, 12, 1145-1166.
- Fry, G. Depression of the activity aroused by a flash of light by applying a second flash immediately afterwards to adjacent areas of the retina. *American Journal of Physiology*, 1934, 108, 701-707.
- Fukada, Y. Receptive field organization of cat optic nerve fibers with special reference to conduction velocity. *Vision Research*, 1971, 11, 209-226.
- Graham, N. Spatial frequency channels in human vision: Detecting edges without edge detectors. In C. Harris (Ed.), *Visual coding and adaptability*. Hillsdale, N. J.: Erlbaum, in press.
- Grüsser-Cornehl, U. Response of movement-detecting neurons of the frog's retina to moving patterns under stroboscopic illumination. *Pflügers Archiv für die gesamte Physiologie des Menschen und der Tiere*, 1968, 303, 1-13.
- Grüsser-Cornehl, U., Grüsser, O.-J., & Bullock, T. Unit responses in the frog's tectum to moving and nonmoving visual stimuli. *Science*, 1963, 141, 820-822.
- Hoffmann, K.-P. Conduction velocity in pathways from retina to superior colliculus in the cat: A correlation with receptive field properties. *Journal of Neurophysiology*, 1973, 36, 409-424.
- Hubel, D., & Wiesel, T. Receptive fields and functional architecture of monkey striate cortex. *Journal of Physiology*, 1968, 195, 215-243.
- Hughes, A., & Whitteridge, D. The receptive field and topographical organization of goat retinal ganglion cells. *Vision Research*, 1973, 13, 1101-1114.
- Ikeda, H., & Obonai, T. Figural aftereffects, retroactive effect and simultaneous illusion. *Japa-*

- nese *Journal of Psychology*, 1955, 26, 235-246. (English summary pp. 290-291.)
- Kahneman, D. An onset-onset law for one case of apparent motion and metacontrast. *Perception & Psychophysics*, 1967, 2, 577-584.
- Kahneman, D. Method, findings, and theory in studies of visual masking. *Psychological Bulletin*, 1968, 70, 404-425.
- Kolers, P. Intensity and contour effects in visual masking. *Vision Research*, 1962, 2, 277-294.
- Kolers, P. *Aspects of Motion Perception*. New York: Pergamon Press, 1972.
- Kolers, P., & Rosner, B. On visual masking (metacontrast): Dichoptic observation. *American Journal of Psychology*, 1960, 73, 2-21.
- Lefton, L. Metacontrast: A review. *Perception & Psychophysics*, 1972, 13, 161-171.
- Matin, E. Light adaptation and the dynamics of induced tilt. *Vision Research*, 1974, 14, 255-265. (a)
- Matin, E. Saccadic suppression: A review and an analysis. *Psychological Bulletin*, 1974, 81, 899-917. (b)
- Matin, E. Saccadic suppression and the stable world. In R. Monty & J. Senders (Eds.), *Eye movements and psychological processes*. Hillsdale, N. J.: Erlbaum, in press.
- Mayzner, M., Tresselt, M., & Helfer, M. A provisional model of visual information processing with sequential inputs. *Psychonomic Monograph Supplements*, 1967, 2 (7, Whole No. 23).
- McCann, G. The fundamental mechanism of motion detection in the insect visual system. *Kybernetik*, 1973, 12, 64-73.
- Noda, H. Sustained and transient discharges of retinal ganglion cells during spontaneous eye movements of cat. *Brain Research*, 1975, 84, 515-529.
- Noda, H., & Adey, R. Excitability changes in cat lateral geniculate cells during saccadic eye movements. *Science*, 1974, 183, 543-545.
- Noda, H., Freeman, R., & Creutzfeldt, O. Neuronal correlates of eye movements in the visual cortex of the cat. *Science*, 1972, 175, 661-664.
- Noda, H., Freeman, R., Gies, B., & Creutzfeldt, O. Neuronal responses in the visual cortex of awake cats to stationary and moving targets. *Experimental Brain Research*, 1971, 12, 389-405.
- Oyster, C., Takahashi, E., & Collewijn, H. Direction-selective retinal ganglion cells and control of optokinetic nystagmus in the rabbit. *Vision Research*, 1972, 12, 183-193.
- Raab, D. Backward masking. *Psychological Bulletin*, 1963, 60, 118-129.
- Reichardt, W. Autocorrelation, a principle for the evaluation of sensory information by the central nervous system. In W. Rosenblith (Ed.), *Sensory communication*. Cambridge, Mass.: M.I.T. Press, 1961.
- Reichardt, W. Movement perception in insects. In W. Reichardt (Ed.), *Processing of optical data by organisms and by machines. Proceedings of the International School of Physics (Enrico Fermi)*. New York: Academic Press, 1969.
- Riggs, L., Merton, P., & Morton, H. Suppression of visual phosphenes during saccadic eye movements. *Vision Research*, 1974, 14, 997-1011.
- Schiller, P., & Smith, M. Detection in metacontrast. *Journal of Experimental Psychology*, 1966, 71, 32-39.
- Schiller, P., & Smith, M. Monoptic and dichoptic metacontrast. *Perception & Psychophysics*, 1968, 3, 237-239.
- Singer, W., & Bedworth, N. Inhibitory interaction between X and Y units in the cat lateral geniculate nucleus. *Brain Research*, 1973, 49, 291-307.
- Sperling, G. Temporal and spatial visual masking. I. Masking by impulse flashes. *Journal of the Optical Society of America*, 1965, 55, 541-559.
- Stigler, R. Chronophotische Studien über den Umgebungskontrast. *Pflügers Archiv für die gesamte Physiologie des Menschen und der Tiere*, 1910, 134, 365-435.
- Stone, J., & Hoffmann, K-P. Very slow conducting ganglion cells in the cat's retina: A major new functional type? *Brain Research*, 1972, 43, 610-616.
- Toch, H. The perceptual elaboration of stroboscopic presentations. *American Journal of Psychology*, 1956, 69, 345-358.
- Weinstein, N. A Rashevsky-Landahl neural net: Simulation of metacontrast. *Psychological Review*, 1968, 75, 494-521.
- Weinstein, N. Metacontrast. In D. Jameson & L. Hurvich (Eds.), *Handbook of sensory physiology* (Vol. 7, part 4: *Visual psychophysics*). Berlin: Springer-Verlag, 1972.
- Weinstein, N., & Growney, R. Apparent movement and metacontrast: A note on Kahneman's formulation. *Perception & Psychophysics*, 1969, 5, 321-328.
- Weinstein, N., Ozog, G., & Szoc, R. A comparison and elaboration of two models of metacontrast. *Psychological Review*, 1975, 82, 325-343.
- Werner, H. Studies on contour: I. Qualitative analyses. *American Journal of Psychology*, 1935, 47, 40-64.
- Wurtz, R. Visual cortex neurons: Response to stimuli during rapid eye movements. *Science*, 1968, 162, 1148-1150.
- Wurtz, R. Comparison of effects of eye movements and stimulus movements on striate cortex neurons of the monkey. *Journal of Neurophysiology*, 1969, 32, 987-994. (a)
- Wurtz, R. Response of striate cortex neurons to stimuli during rapid eye movements in the monkey. *Journal of Neurophysiology*, 1969, 32, 975-986. (b)
- Wurtz, R. Visual receptive fields of striate cortex neurons in awake monkeys. *Journal of Neurophysiology*, 1969, 32, 727-742. (c)
- Wurtz, R., & Goldberg, M. The role of the superior colliculus in visually-evoked eye movements. *Bibliotheca Ophthalmologica*, 1972, 82, 149-158.

On the Role of Context Effects in Psychophysical Judgment

Norman H. Anderson
University of California, San Diego

Functional measurement provides a general perspective on the role of context effects in judgment theory. Instead of being treated as undesirable biases, context effects can be incorporated into substantive theory and thereby made a base and frame for measurement. The first part of the article illustrates this approach in detail by applying functional measurement and Parducci's range-frequency theory to the analysis of an iterative technique for category ratings propounded by Stevens. The iterative technique is not disconfirmable and so not meaningful by itself. Range-frequency theory is shown to provide a theoretical foundation for the iterative technique that is testable. It also provides a basis for interpreting the limit function from the iterative technique as the psychophysical law. This analysis is extended to handle a very general range-frequency model. The second part of the article extends this perspective to context effects in general. Among other things, the functional measurement diagram is used to show that in contrast to the traditional approach based on magnitude estimation, functional measurement (a) allows context effects to be treated as substantive phenomena, interesting in themselves, useful in scaling, not just undesirable biases; (b) provides a clear conceptual definition of interval scale; (c) does not require sensation to be conscious; and (d) derives the psychophysical law from the more basic psychological law.

Judgment of any one stimulus is often affected by other, contextual stimuli. Some of these context effects, those involved in the size-weight illusion for example, make good ecological sense. Others seem to have little rational or ecological relation to the judgment and are typically treated as biases.

One such bias has attracted special attention in psychophysical judgment, namely, the distribution effect in judgments of single stimuli. The same stimulus may be rated differently when it is a member of a normal-type distribution, for example, than when it is a member of a flat, rectangular distribution, even though both distributions cover the same range of stimuli. Although at-

tempts have been made to treat such distribution effects in terms of contrast and assimilation, that is, as changes in perceptual value proper, the more common view is that they merely reflect the way in which the response language is used. If the overt response is to be treated as a direct measure of perceptual magnitude, then such context effects are clearly undesirable biases.

These distribution effects, which can be quite large, have been emphasized by Stevens (1957, 1971) as one of the main faults of category scaling. However, Stevens has also suggested an iterative technique that is intended to converge on a limiting distribution of stimuli in which the bias is "neutralized" (Stevens & Galanter, 1957, pp. 381-382).

No theoretical analysis of the iterative technique has been given. About all that is known is that it seems to converge rapidly in practice. The most important issue centers on the interpretation of the response function produced by the technique. Stevens and Galanter (1957, p. 382) state that it will be the "pure form of the category scale" but "not in general an equal-interval scale of

This research was supported by National Science Foundation grants BMS 74-19124 and GS-36918 and by grants from the National Institute of Mental Health to the Center for Human Information Processing, University of California, San Diego. I wish to thank Edward Calvin Carterette, Hannes Eisler, Stephen Milne, Henry Montgomery, Allen Parducci, and David J. Weiss for their helpful comments on a draft of this paper.

Inquiries concerning this article should be addressed to Norman H. Anderson, Department of Psychology, C-009, University of California, La Jolla, California 92093.

the psychological magnitude." However, Pollack (1965, p. 565) speaks of the potential of the iterative technique for producing "unbiased psychophysical functions," and Montgomery and Eisler (1974, p. 441) treat the result as an "equal interval scale." To clear up these discrepant interpretations, it is necessary to determine what is assumed when using the iterative technique and what it does.

The present article has three interrelated goals. The first is to provide a theoretical foundation for Stevens' iterative scaling procedure. That will be done using a mixture of ideas from integration theory and functional measurement (Anderson, 1974a, 1974d) and from Parducci's (1965, 1974) range-frequency theory of category judgment.

The second goal is to contribute to the general theory of category judgments that stems from Parducci's range-frequency theory. These first two goals are the subject of the first part of this article.

The third goal is to present a general perspective on context effects. Instead of being treated as undesirable biases, context effects can be incorporated into substantive theory. This perspective, illustrated by the analysis of the distribution effect, is seen to have wide relevance in the second part of this article. By taking a substantive attitude toward context effects, functional measurement is able to provide a unified approach to many traditional problems in psychophysics.

THEORETICAL ANALYSIS

Notation

The notation for the subsequent theoretical analysis is simple, but it will be helpful to summarize it here. In the experimental task under consideration, the subject is presented one stimulus at a time and estimates the magnitude of some quality of that stimulus, such as the loudness or pitch of a tone. The set or distribution of stimuli presented to the subject is assumed to have a physical metric. For convenience, the same symbol S will be used to denote the stimulus and its physical value. These values are taken to lie between

S_L and S_U , the lower and upper bounds of the stimulus distribution.

Three functions of S will be needed. The first is the response function $R(S)$; it denotes the theoretical or the mean response to S . The second is the cumulative frequency function $F(S)$; it can be treated as the percentile rank of S . The third and last is the psychophysical function $V^*(S)$; it represents the perceptual magnitude of S .

These three functions will be represented in a common, normalized, percentile form. Thus, $F(S)$ will be measured as the ordinary frequency percentile, with values between 0 and 100. Correspondingly, the values of $R(S)$ will arbitrarily be taken to lie on a 0-100 scale. To represent the psychophysical function in an analogous percentile form requires a linear transformation on $V^*(S)$ to obtain the percentile form $V(S)$:

$$V(S) = 100[V^*(S) - V^*(S_L)] / [V^*(S_U) - V^*(S_L)]. \quad (1)$$

In this expression the denominator is the subjective stimulus range, and the numerator is the perceptual value of S relative to a zero point at the lower end of the range. Thus defined, $V(S)$ is a linear transformation of $V^*(S)$ and may also, for present purposes, be treated as the psychophysical function.

For expositional purposes, it will often be convenient to speak in terms of discrete stimulus distributions and discrete response categories. In the theoretical analyses, however, it will be assumed that the response scale is effectively continuous, a graphic rating scale, for example. Also, the distribution of stimuli presented for judgment will be assumed to be effectively continuous so that, in particular, $R(S)$ is defined for all S . These continuous distributions can be viewed as approximations to the histogram forms of the discrete distributions, or vice versa. In the iterative technique itself, each stimulus is ordinarily presented equally often, so each histogram component would have constant area, varying in height to reflect the interstimulus distances.

Subscript n will denote the n th iteration, and subscript ∞ will denote the limit of the

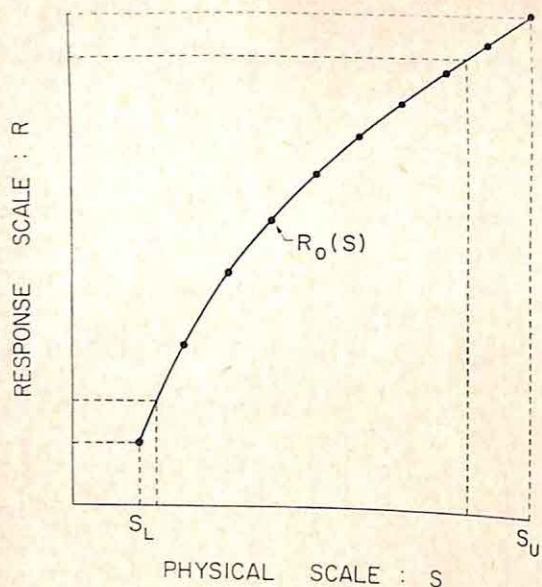


FIGURE 1. Illustration of the operation of the iterative technique (S_L = lower bound of stimulus distribution; S_U = upper bound).

iterative process. Subscripts i , j , and k will denote arbitrary distributions of stimuli.

Iterative Technique

The iterative technique is illustrated in Figure 1, patterned after Pollack (1965). The vertical axis represents the rating response, and the horizontal axis represents the stimulus metric, such as db., g, or cm.

The curve $R_0(S)$ represents the response to an arbitrary set or distribution of stimuli, each presented with equal frequency. This response function is used to select a new distribution for the iteration. The working assumption is that the subject tends to use the response categories equally often, and the goal is to arrive at a final distribution in which that tendency is "neutralized."¹

¹ Discussions of the iterative technique typically speak of the subject's "expectations" that the response categories will be used equally often, but remain vague about how those expectations get translated into responses. Accordingly, the present exposition has adopted Parducci's conception of response tendencies associated with each category. These response tendencies, which are one determinant of the actual response, are assumed to depend only on the response scale itself, independent of the actual stimuli. This conception, given in more detail below, seems to capture the intent of the iterative technique.

In Figure 1, the initial distribution of stimuli was chosen to be equally spaced on the physical scale. The response curve $R_0(S)$ is bowed downward, high responses being more frequent than low responses. To equalize the frequency of response usage, it is necessary to increase the density of stimuli in the lower part of the range where the response curve rises faster. That is done as follows.

Begin by dividing the vertical range between the smallest and largest responses into k equal parts. From each dividing point, project a horizontal to the curve $R_0(S)$, then drop a vertical to the stimulus axis. The set of $k + 1$ stimuli so obtained, including the two end stimuli, constitutes the new distribution, each stimulus being presented equally often. This procedure is illustrated by the dashed curves in Figure 1 for the first two and last two stimuli. As can be seen, the first two stimuli are considerably closer in the new than in the old distribution; the last two stimuli are considerably farther apart.

This unequal stimulus spacing will tend to equalize the frequency of response usage. This new distribution is presented for judgment and yields a new response curve, $R_1(S)$, which is treated similarly as the basis for the next iteration. This process is to be repeated until it converges.

This description of the iterative technique shows its essential content to be that the response function on one iteration defines the frequency function for the next iteration. In present notation,

$$F_n(S) = R_{n-1}(S). \quad (2)$$

This equation mirrors the description of the iterative technique in Figure 1. By this technique, the frequency of stimuli that lie between S_1 and S_2 in the new distribution is just $R_{n-1}(S_2) - R_{n-1}(S_1)$.

If the iterative technique converges, then in the limit the frequency function becomes identical to the response function:

$$F_\infty(S) = R_\infty(S). \quad (3)$$

Thus, a fixed increment on the response scale corresponds to an equivalent increment in

the frequency function.² Note that this result does not define the shape of the response function; that depends on the physical metric of S .

The rationale leading to Equation 3 is essentially a formalization of the iterative technique as it has been presented in the literature. This formalization makes clear the aim and nature of the iterative technique, but it also brings out some serious limitations. First is the lack of any proof that the iterative technique converges. The previous reliance on experimental demonstrations of convergence is sensible in itself, but

² A curious aspect of the iterative technique is that the response categories will not in general be used equally often in the limit, contrary to assertions by Stevens (1957) and Stevens and Galanter (1957). They give no basis for this assertion, but presumably it was thought to be required by the claims made for the iterative technique. If indeed the iterative technique does produce a stimulus distribution that neutralizes potential bias caused by the tendency to use the response categories equally often, then it would seem that such neutralization would require equal category usage. However, that will be true only under special conditions.

If response variability was zero, then each response would be used equally often in the limit. Equation 3 implies that a unit increment in stimulus frequency corresponds to a unit increment on the response magnitude scale, and so also, given zero response variability, to a unit increment in response frequency.

However, if response variability is not zero, then the responses that actually fall within any interval on the response scale will include random strays from adjacent intervals. Unless the responses that stray out of each interval just cancel those that stray in, which is not generally to be expected, the response categories will not be used with equal frequency.

This point was first made by Montgomery and Eisler (1974), who also seem to be the first to report relevant data. In fact, their response frequencies had a U-shape, with more responses in the two end categories and fewer in the middle. They note that the response variability was very low for the two end stimuli, which were, accordingly, almost always assigned to the two end response categories, very seldom to the adjacent categories. However, response variability was markedly higher for the adjacent stimuli, which were assigned to the two end response categories with appreciable frequency. Thus the two end categories gained in frequency of usage at the expense of the adjacent categories. The loss suffered by the latter would in turn propagate inward to the middle categories, augmented perhaps by further unequal response variability.

that leaves the equal-frequency assumption untested and its role unclear. This assumption seems crucial, yet the iterative technique seems rather to impose it than to test it. If convergence can occur with unequal-frequency tendencies, then the iterative technique may not only fail to neutralize the one bias, but also may introduce another bias of its own.

No less serious is the fact that the iterative technique provides no substantive interpretation of the limit response function $R_\infty(S)$. The disagreement among the three quotations above reflects this shortcoming.

Both limitations of the iterative technique stem from the lack of any underlying theory. To give meaning to the iterative technique, therefore, it is necessary to have recourse to some substantive theory. Parducci's range-frequency theory, which will be employed here, has the great advantage that it can be generalized to provide an analytical treatment even when the equal-frequency assumption fails.

Range-Frequency Theory

In Parducci's (1965, 1974) theory of category judgment, the overt response is a compromise between two covert response tendencies or values. These are the range value and the frequency value. The range value can be represented as the range percentile $V(S)$ of Equation 1; it is just the perceptual magnitude of S , expressed in terms of its relative location on the subjective scale between the two end stimuli. For the time being, it is assumed that the subject tends to use each response category equally often, as with the iterative technique above. Accordingly, the frequency value will be represented in cumulative form as the ordinary frequency percentile. This definition of the frequency value exactly parallels that of the range value.

In general, the range tendency and the frequency tendency will disagree about the response to any given stimulus. Responses controlled by the equal-frequency tendency alone would produce a uniform response distribution, each response being used equally often regardless of the actual distri-

bution of stimulus values. Responses controlled by the range tendency alone would exactly mirror the (subjective) stimulus distribution.

The actual response to any given stimulus is assumed to be a compromise between the range tendency and the frequency tendency. Mathematically, this range-frequency compromise is assumed to follow an averaging model:

$$R(S) = wV(S) + (1 - w)F(S) + \epsilon(S), \quad (4)$$

where w is a relative weight, $0 < w < 1$, and $\epsilon(S)$ represents response variability. The response variability is assumed to have zero mean for each S and can be ignored in what follows. Accordingly, $R(S)$ will refer to the mean value of the response.

A basic assumption in range-frequency theory is that the range value is independent of the stimulus distribution. This assumption does not disallow contrast effects, however, because the zero of the perceptual scale is lost in the percentile representation of Equation 1. Thus, range-frequency theory is not inconsistent with an interpretation of distribution effects as changes in perceptual value proper, though it does require $V^*(S)$ itself to be linearly related across different distributions.

It should be emphasized that the present interpretation of the range value as the basic perceptual stimulus value is an assumption. Parducci (1974, p. 137) prefers a nominal interpretation as a model parameter, in part because of a proper concern that the range value must yet be shown to have cross-task generality before it can be considered as a basic scale value. Applications of functional measurement in other tasks (e.g., Anderson, 1972a, 1974c; Weiss, 1975) have supported cross-task invariance, and have illustrated how invariance can provide a powerful tool for analyzing the structure of perception and judgment. This matter deserves systematic study within range-frequency theory, especially because of its bearing on the nature of the rating process.

Range-Frequency Analysis of the Iterative Technique

The range-frequency model of Equation 4 can be applied directly to represent the iterative technique:

$$R_n(S) = wV(S) + (1 - w)F_n(S). \quad (5)$$

The iteration index n is omitted from $V(S)$ since it is assumed to be independent of the stimulus distribution. From Equations 5 and 2,

$$R_n(S) = wV(S) + (1 - w)R_{n-1}(S). \quad (6)$$

This difference equation converges under quite general mathematical conditions. Formally, the solution can be obtained by setting $R_n(S) = R_{n-1}(S) = R_\infty(S)$, and solving to obtain

$$R_\infty(S) = V(S). \quad (7)$$

Equation 7 says that the limit response function is equivalent to the psychophysical function. If range-frequency theory is correct, therefore, then the outcome of the iterative technique has a simple, useful meaning.

The range-frequency model is sufficient to ensure convergence, but other models will do the same. An interesting case is the generalization of the range-frequency model to allow for differential weighting. Thus, the w parameter in Equation 4 would not be constant, but instead would depend on S . For stimuli near the ends of the distribution, for example, the perceptual term $V(S)$ might have greater weight because these stimuli are more discriminable than stimuli in the middle of the distribution.

The iterative technique will converge even with differential weighting. Furthermore, Equation 4 still leads to Equation 7, so that the limit response function is equivalent to the psychophysical function even in this generalized case. This result is potentially important because it means that the iterative technique allows for other biases than those based on frequencies of category usage. However, this result rests on the equal-frequency assumption. If that fails, then the limit function no longer is meaningful.

Comments on Assumptions

Basic to the above analysis is the equal-frequency assumption, that all responses tend to be used equally often. This assumption will not be correct in general. As an extreme example, one would not ordinarily expect the two response categories, *tiny* and *larger than tiny*, to be used equally often in judgments of size.

But even for an ordinary rating scale, with categories ranging from, say, *very very small* to *very very large*, a reasonable expectation would be for a normal-like distribution of usage. The normal distribution has ample basis in daily experience, in which the rough rule is that most events are near average in magnitude and extreme events are infrequent. In contrast, the flat, equal-frequency distribution is rare in daily experience.

Clearly, the equal-frequency assumption demands special attention. It may turn out to be correct, but it requires better empirical support than is currently available.

A second basic assumption is that the range-frequency model represented by Equation 4 is correct. In support of this assumption, Parducci's theory has had substantial success. The range-frequency model has been dramatically superior to Helson's (1964) adaptation-level theory of category judgments in various qualitative tests (Parducci & Perrett, 1971). Although the published quantitative tests present a cloudy picture, the model does have considerable potential.

There is a final basic assumption, one that has a somewhat different conceptual status and function. This is the assumption that the overt response is an interval scale. But the iterative technique will converge even with an ordinal response scale; the assumption that the overt response is an interval scale serves only the one purpose of interpreting the limit response function as the psychophysical function. This assumption is not accepted by all, as shown by the quoted passages above, and its nature and status require a more detailed discussion. That discussion is deferred until the second part of this paper.

Generalized Range-Frequency Model

The equal-frequency assumption is basic to the iterative technique, and it has been standard in nearly all applications of range-frequency theory. However, there is reason to suspect the equal-frequency assumption and rather little solid evidence to support it. Accordingly, it is desirable to reconsider the above analyses while allowing for tendencies to use the response categories with unequal frequency. This section considers the generalization of the range-frequency model. The next section shows how the iterative technique suffers under the unequal-frequency assumption. The following section will present exact tests of the range-frequency model.

To handle the unequal-frequency assumption, it is useful to begin by stating a very general version of the range-frequency model in which the overt response is an average of a generalized range value $\tilde{V}(S)$ and a generalized frequency value $\tilde{F}(S)$. Thus,

$$R(S) = w \tilde{V}(S) + (1 - w) \tilde{F}(S). \quad (8)$$

No assumption is made about the form of $\tilde{V}$ or $\tilde{F}$ except that they be expressed in the percentile form indicated above.

The essential properties of this generalized range-frequency model are additivity and independence. The generalized range value is assumed to be independent of the distribution of stimuli, and the generalized frequency value is assumed to be independent of the stimulus magnitudes. These properties alone are sufficient to provide a test of the model, and that is illustrated in Array 17 below. Here it is necessary to consider how the unequal-frequency assumption affects the specification of $\tilde{V}$ and $\tilde{F}$.

Parducci (1965, 1974) allows in principle for tendencies to use the response categories with unequal frequencies. His specification of $\tilde{F}(S)$ seems straightforward. The frequency scale is to be divided into frequency categories, one for each response category, of size equal to the magnitude of the tendency to use the corresponding response category. The generalized frequency value of any stimulus then corresponds to the value

of the corresponding response category. For the continuous case considered here, let $G(R)$ denote the given response tendencies, in cumulative percentile form. This G function then also corresponds to the division on the frequency scale. Accordingly, its inverse function, denoted by E , defines the response value, that is, the generalized frequency value of stimulus S :

$$\tilde{F}(S) = E[F(S)]. \quad (9)$$

According to Parducci (1974, p. 136), the range value of S depends only on the two end stimuli and on S itself. That implies that the range value is independent of how the categories are defined, an assumption that has also been used by Birnbaum (1974). Accordingly,

$$\tilde{V}(S) = V(S). \quad (10)$$

This assumption that the range value is independent of the response scale seems open to question. For example, the two two-category response scales, *tiny-larger than* and *small-large*, seem to imply different cutting points on the stimulus scale, regardless of any frequency tendencies associated with the categories per se.

On this basis, an alternative specification of the range value may be suggested that parallels the above specification of the frequency value. Thus, the perceptual scale $\tilde{V}(S)$ would be divided into unequal-sized categories in exactly the same way as the frequency scale. In that case,

$$\tilde{V}(S) = E[V(S)]. \quad (11)$$

A simple test of Parducci's specification of $\tilde{V}(S)$ can be obtained by using the method of range extension described in the next section. If $\tilde{V}(S) = V(S)$, as Parducci assumed, then Equation 15 below can be generalized to handle the unequal-frequency assumption simply by replacing $F(S)$ with $\tilde{F}(S)$. It follows that the range values of each common stimulus are proportional and nearly related across the several distributions. However, this relation is nonlinear $\tilde{V}(S) \neq V(S)$, as in Equation 11. The diagnostic power of this test, of course, de-

pends on obtaining substantial deviations from the equal-frequency assumption.

This problem of specifying the generalized range value cannot be pursued further here. However, it does bear on the interpretation of the range value as a perceptual magnitude. Also it is of interest because of its close connection with the problem of response emission. In general, it seems clear that the specification of $\tilde{V}(S)$ rests on some assumption, explicit or implicit, about the response metric. Perhaps Equation 11 applies to the verbal category scales that have been typical of range-frequency experiments, and Equation 10 would hold for a numerical or graphic rating scale. In the latter case there are no response categories per se, and a direct correspondence could be set up between the perceptual magnitude and the response frequency tendencies that might also be operative.

Iterative Technique Under the Unequal-Frequency Assumption

If the equal-frequency assumption is invalid, then direct application of the iterative technique will not in general yield a meaningful result. To illustrate this, suppose that $\tilde{V}(S) = E[V(S)]$. Then the range-frequency generalization of Equation 6 is

$$R_n(S) = wE[V(S)] + (1-w)E[R_{n-1}(S)].$$

If $V(S)$ was a solution of this equation, so that $R_\infty(S) = V(S)$, then this equation would require

$$V(S) = E[V(S)].$$

Since E is nonlinear by assumption, $V(S)$ cannot be a solution. The same conclusion holds generally, except for unusual definitions of $\tilde{V}(S)$. If the equal-frequency assumption is invalid, therefore, the iterative technique will not yield the psychophysical function.

Goodness of Fit: Range-Frequency Model

Since the range-frequency model applies to any stimulus distribution, it can be tested without employing the iterative technique.

Indeed, the strongest tests are obtained by using specially chosen distributions, and some striking qualitative results have been obtained in this way. Thus, Parducci and Perrett (1971; see Anderson, 1974a, Figure 6) predicted and obtained a crossover of the response junctions obtained from a U-shaped and a normal distribution. In the same way, Birnbaum (1974) predicted and obtained a double crossover of the response functions from two bimodal distributions. Other theories of category judgment, such as that of Helson (1964), cannot account for these crossovers.

Quantitative tests are also necessary, especially if the range-frequency model is to be applied to psychophysical scaling. Unfortunately, the quantitative tests have not been generally satisfactory. The present paper takes a functional measurement approach in which the range-frequency model itself becomes the base and frame for scaling.

Two main methods of testing are possible: the constant range method, and the range extension method. Both rely on comparing the response to the same stimulus across different distributions. The constant range method was first used by Parducci and Perrett (1971). The range extension method was introduced by Anderson (1974a; Note 1, Section II.B.8), who gave a functional measurement analysis of both methods. In addition, a power series method, noted below, has been developed by Birnbaum (1974), who has also stressed the functional measurement approach to the study of context effects in psychophysics.

The essential ideas of the constant range method and the range extension method will be developed more fully here, although no attempt will be made to cover all the various statistical considerations. Each method will be considered first under the equal-frequency assumption, and then they will be combined to handle the generalized range-frequency model.

The first method compares across stimulus distributions that have the same end stimuli plus one or more additional common stimuli. By virtue of Equation 1, each such common stimulus has the same range per-

centile in each distribution. For the i th distribution, therefore,

$$R_i(S) = wV(S) + (1-w)F_i(S). \quad (12)$$

Thus, the observable $R_i(S)$ is a linear function of the known value of $F_i(S)$ for fixed S . Standard regression analysis can be used to test this linearity prediction with only one stimulus common to three distributions. A more efficient statistical test can be obtained by pooling data across distributions with different pairwise common stimuli.

A striking implication of Equation 12 is that just one such test can determine the complete psychophysical function $V(S)$. Both the response function and the frequency function are known, and the w parameter is estimated by the regression analysis. Accordingly, Equation 12 yields an estimate of $V(S)$ in terms of known quantities:

$$V(S) = (1/w) [R_i(S) - (1-w)F_i(S)]. \quad (13)$$

Of course, each distribution should yield the same $V(S)$, and that provides a further check.

The main idea of this constant range method was originally used on a post hoc basis by Parducci and Perrett (1971, pp. 445-447) with the arbitrary assumption that $w = 1/2$. The model did not do well in these tests, but no information is given that suggests the nature or cause of the discrepancies.

The second, range-extension method compares across stimulus distributions that have different ranges. To illustrate the essential idea, consider several distributions that are identical over their lower domain, from S_L up to some medium stimulus S_M . These distributions are assumed to differ above S_M by having, in particular, different upper end stimuli, S_{U_i} , for distribution i .

Any stimulus in the lower domain necessarily has the same frequency percentile in all distributions because they are identical over the lower domain. Furthermore, the range percentiles of these stimuli are linearly related by virtue of Equation 1. Thus,

the range percentiles for distributions 1 and i obey the relation,

$$V_i(S) = c_i V_1(S), \quad (14)$$

where c_i is the ratio of the subjective ranges of the two distributions, namely, $[V^*(S_{U1}) - V^*(S_L)]/[V^*(S_{Ui}) - V^*(S_L)]$.

Under these conditions, the range-frequency model of Equation 4 becomes

$$R_i(S) = wc_i V_1(S) + (1 - w)F(S). \quad (15)$$

Since $F(S)$ is the same for common stimuli in the lower domain, subtraction among distributions i , j , and k yields

$$\begin{aligned} R_i(S) - R_j(S) &= (c_i - c_j)wV_1(S), \\ R_i(S) - R_k(S) &= (c_i - c_k)wV_1(S). \end{aligned} \quad (16)$$

Accordingly, the two sets of differences on the left should plot as straight-line functions of each other, because both are proportional to $V_1(S)$. Further, if this test is satisfied, then the differences provide an estimate of the complete psychophysical function $V_1(S)$. This test is not statistically optimal, but it does allow a quick inspection of the data.

A formal and more efficient statistical test can be obtained by viewing Equation 15 as a compound multiplying-adding model. The bilinear test for multiplying models does not apply directly, but the model still has a semi-linear form that allows direct application of analysis of variance (Mandel, 1964, Section 9.4). In this test, the common stimuli in the lower domain constitute one factor of the design, and the several distributions constitute the other factor. This test also holds with $F(S)$ replaced by $\tilde{F}(S)$, as is suggested by Equation 16, in which the frequency value has cancelled out.

For the generalized range-frequency model, an exact test can be obtained by combining the two test methods discussed above. To illustrate, consider two distributions with the same end stimuli. Choose a larger upper end stimulus and construct two new distributions by stretching out the two original distributions in such a way that they remain unchanged below some intermediate value S_M . Any common stimulus less than S_M

will then have the same frequency percentile in either distribution and its extension; it will also have the same range percentile in each pair of distributions with the same end stimulus. Thus, the theoretical values can be expressed in a 2×2 array:

$$\begin{array}{cc} w\tilde{V}_1(S) + (1 - w)\tilde{F}_1(S) & w\tilde{V}_2(S) + (1 - w)\tilde{F}_1(S) \\ w\tilde{V}_1(S) + (1 - w)\tilde{F}_2(S) & w\tilde{V}_2(S) + (1 - w)\tilde{F}_2(S). \end{array} \quad (17)$$

Here $\tilde{F}_1$ and $\tilde{F}_2$ represent the generalized frequency values defined by the two original distributions over the lower range of stimuli. Analogously, $\tilde{V}_1$ and $\tilde{V}_2$ represent the generalized range values defined by the original range and the extended range, respectively.

The key property of the 2×2 array is that the interaction term is zero. This follows immediately because the differences between the two entries in each column are equal. Graphically, the data should exhibit parallelism. Although the model can be tested using only a 2×2 design, useful parameter estimation would, of course, require more than two levels.

This test can obviously be extended to more than two original distributions and to more than two upper end stimuli, combined in a two-way factorial design. The data from such a design should plot as a set of parallel curves. This parallelism test has the attractive feature that no specific parameter estimation is required. Indeed, as is explicit in Array 17, no assumption is needed even about the form of the generalized range and frequency functions.

The above tests all have certain desirable statistical properties. In particular, they are exact. That is, they do not rest on approximations to an unknown function, and they do not use external parameter estimates of unknown reliability and/or validity.

A related application of functional measurement to range-frequency theory has been given by Birnbaum (1974). An important outcome of this work, in which subjects gave verbal category ratings of magnitudes of three-digit numbers, was the double cross-over in the response functions produced by

two different stimulus distributions. As noted above, this result is contrary to Helson's (1964) theory of adaptation level but agrees with the prediction from range-frequency theory.

Birnbaum also gave a quantitative treatment of the equal-frequency model of Equation 4, expressed in somewhat different notation. An interesting feature of his analysis was that $V(S)$ in Equation 4 was replaced by a power series in the objective number metric of S . The power series approach has the special advantage that it requires only the data from a single distribution to get a test of fit plus an estimate of the psychophysical function. The test of fit, however, is not exact. To establish the basic model, therefore, it would seem preferable to use one of the above exact tests that requires no assumption about the form of the psychophysical function. Once the basic model has been established, then power series and other approximate methods can be very useful for estimation.

One final question remains to be considered. That is the assumption that the overt response is an interval measure. Within the range-frequency model, response nonlinearity is conceptually distinct from the unequal-frequency assumption, and even the general model of Equation 8 can fail because of nonlinearities in the overt response. In principle, response nonlinearity can be rectified by monotone transformation. Thus, Birnbaum, Parducci, and Gifford (1971) found discrepancies from the model, which they attributed to nonlinear response bias, and which they attempted to remove by monotone response transformation. Their analysis was based on the equal-frequency assumption, however, and possibly that assumption was at fault rather than the response scale.

The general problem of response transformation can, in principle, be handled by using functional measurement in connection with larger versions of Array 17. If the parallelism test succeeds, that supports both the generalized model and the interval property of the response measure. If the parallelism test fails owing only to nonlinearity

in the response measure, then parallelism can be restored by monotone rescaling (Anderson, 1974a, Section II.A.6).

To sum up, functional measurement provides an effective analysis of the range-frequency model under the most general conditions. It is not necessary to make the equal-frequency assumption. Nothing need be assumed about the psychophysical function. Neither is it necessary to have a linear or equal-interval response scale. All that is at issue is the basic algebraic structure of the model itself; that constitutes the base and frame for the functional measurement analysis.

Notes on Experimental Design

There are various points of design and procedure that should be kept in mind in experimental applications of the above results. One of the foremost relates to the assumption of continuity in the response scale that was used in the theoretical analysis. In practice, most investigators have used a small number of categories, and this discreteness may introduce biases that are not represented in the model. Although this problem may not be overly serious, the use of a continuous graphic rating scale has much to recommend it. Not only would it essentially bypass this discreteness problem, but it would also help eliminate sequential effects from previous trials on the response to the present stimulus (Anderson, 1969, 1971, 1972b; Parducci, 1974).

Since the iterative technique depends heavily on the equal-frequency assumption, it would seem desirable to facilitate it by instructions. The instructions that have been used often seem likely to facilitate unequal-frequency tendencies. Of course, such instructions would be appropriate for studying the unequal-frequency assumption itself.

The data analysis is much simpler if the rating response is an interval scale. The experimental work on information-integration theory has used rating scales with good success, and this work suggests that only a few modest precautions are necessary (Anderson, 1974a, Section II.A.7; 1974d, Section 2.1).

One precaution that deserves special mention is the use of end anchors, special stimuli that are intended to define and tie down the usage of the rating scale. They are usually chosen to be just noticeably more extreme than the regular experimental stimuli. They are presented in the preliminary practice to help set up the frame of reference and may be repeated during the course of the experiment to keep response usage stable. There is ample (if scattered) evidence for a variety of end effects on both the response and the stimulus sides. The stimulus end anchors offer the promise of fractionating out these end effects from the proper data, and they have other advantages as well.

There is some reason to suspect that the use of stimulus end anchors, together with preliminary practice to set up the stimulus-response frame of reference, can markedly reduce the distribution effect. That could reduce power in the tests of the range-frequency model. However, the use of stimulus end anchors does not seem to violate any assumptions of either the iterative technique or of range-frequency theory, and it is recommended that they be used more widely.

Finally, it is worth noting that much of the above analysis can be adapted for use with verbal or symbolic stimuli. For example, the subject might be asked to judge likableness of persons described by personality-trait adjectives or by photographs. Since such stimuli lack a physical metric, it would be necessary to obtain an auxiliary rank ordering of the stimuli, presumably on a single-subject basis, to calculate the frequency percentiles. These percentiles would be subject to error, which would need to be considered in the statistical analysis (see e.g., Isaac, 1970). Other than that, relations such as Equation 12 would apply as well to symbolic stimuli as to psychophysical stimuli. Equation 13 could then be used to obtain $V(S)$, which, though artificial as a psychophysical function, could provide a convenient method for stimulus scaling.

COMMENTS AND CONCLUSIONS

Iterative Technique

The present analysis helps clarify the nature and usefulness of the iterative tech-

nique. It has two limitations and one advantage. The first limitation is the lack of protection against failure of its basic equal-frequency assumption, that the subject has covert tendencies to use the response categories equally often. The iterative technique is a statistical routine that applies the same regardless of whether the tendencies to use the response categories are equal or unequal. When these tendencies are unequal, however, the iterative technique does not neutralize them but instead injects an additional bias of its own.

The second limitation of the iterative technique is that it provides no basis for interpreting the limit response function even when the equal-frequency assumption is true. Any such interpretation rests on establishing some substantive model such as the range-frequency model. The meaning of the limit function and indeed of the iterative technique itself arises from the underlying substantive model.

Despite these limitations, the iterative technique may have unique potential for scaling. If the equal-frequency assumption is true, then the limit response function will estimate the psychophysical function under more general conditions than are readily amenable to the range-frequency model. In that case, the iterative technique will neutralize biases other than the equal-frequency bias, such as the differential, stimulus-dependent weighting discussed above. If the equal-frequency assumption can be established, either in general or through special instructions and procedure, then the iterative technique may be quite useful.

The present interpretation that the iterative technique yields the psychophysical function rests on an assumption beyond the iterative technique itself. That assumption is that the rating response is an interval scale. But that assumption is not used in the iterative technique, which will converge even though the response is only an ordinal scale. In that case, of course, the limit function will not yield the psychophysical function.

Just that position was taken by Stevens, who considered rating scales to be only ordinal scales on most psychophysical dimen-

FUNCTIONAL MEASUREMENT DIAGRAM

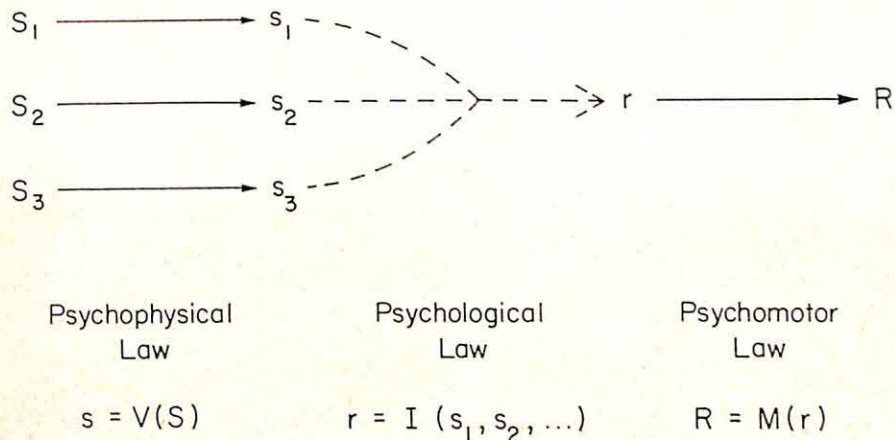


FIGURE 2. Functional measurement diagram. (Physical stimuli S impinge on the organism and are converted by the valuation operation or psychophysical law V into psychological stimuli s . The psychological stimuli are combined according to the integration function or psychological law I to yield an implicit response r . The implicit response is transformed into the overt observable response R according to the output function or psychomotor law M .)

sions. In Stevens' view, even the graphic rating scale is a "partition" scale and so not in general an interval scale. The present interpretation requires justification, therefore, and that will be given in two sections below. Before doing so, it is necessary to clarify what is meant by the term *interval scale*.

Meaning of Interval Scale in Functional Measurement

As used here, the term *interval scale* refers to a linear relation between two quantities. Typically, one quantity is observable and is used as a measure of the other quantity, which is unobservable. Accordingly, the above statement, "the rating response is an interval scale" is ambiguous because only one quantity is explicitly specified. There are two candidates for the other quantity.

The customary approach takes this other quantity to be the subjective magnitude or "sensation," here denoted by $V(S)$. Since distribution effects would in general produce a nonlinear relation between $V(S)$ and the overt response $R(S)$, the overt response would not be an interval scale. To an investigator like Stevens, who aimed at de-

termining the psychophysical function by direct estimation, this nonlinearity between $R(S)$ and $V(S)$ would naturally be considered as an undesirable bias.

A different conceptualization is embodied in functional measurement. The other quantity in the interval scale relation is not $V(S)$, but rather a covert response. This conceptualization is made explicit in Figure 2, which presents a reduced version of the functional measurement diagram (Anderson, 1970b, Figure 1). In this diagram, physical stimuli, denoted by S , impinge on the organism and are processed by the valuation operation V to become psychological stimuli, denoted by s . Thus, $s = V(S)$ defines the psychophysical function. These psychological stimuli are then integrated according to a psychological law I to produce a covert response r . The overt response R is said to be an interval scale when it is linearly related to r .

A vital feature of Figure 2 is the distinction between s and r . The customary approach, based on judgments of single stimuli, aims at using R as a direct measure of s , ignoring r , or assuming r and s to be

linearly related. The functional measurement approach aims at using R as a measure of r .

The necessity of and the justification for the functional measurement approach are well illustrated by the size-weight illusion (Anderson, 1970a). Felt heaviness, which is the subjective magnitude that is accessible to introspection, results from a preconscious integration of information from two stimulus sources, gram weight and visual appearance. A measure of felt heaviness will not provide the psychophysical function for gram weight because the heaviness sensation is also affected by visual appearance. It might seem, therefore, that the visual appearance constitutes a biasing contextual effect that should be eliminated by experimental means.

Nevertheless, it is still a legitimate and important question whether or not the overt response is an interval scale of the conscious sensation of heaviness. In terms of Figure 2, this question reduces to asking whether or not R is a linear function of r .

This question is decidable. The basis of decidability is the psychological law I . If I is an additive function, for example, then the data from a factorial stimulus design should plot as a set of parallel lines. Observed parallelism thus provides joint support for the additive model and for the interval property of the response scale.

Furthermore, this approach can also yield the psychophysical law. For if the additive model is established, then the marginal means of the factorial design estimate the s values on validated interval scales (Anderson, 1970b, 1974a). In the size-weight example, the row means would constitute an interval scale of the subjective value of visual appearance—as it functioned within the given task. Similarly, the column means would constitute an interval scale of the subjective value of gram weight. The graph of these subjective values as a function of the corresponding objective, physical values determines the psychophysical law (Anderson, 1972a, Figure 3; Weiss, 1975, Figure 2).

Rating Scales as Interval Scales

Experimental work on information integration theory has yielded extensive evi-

dence that the ordinary rating method can provide interval response scales. The basic logic has already been set out in the discussion of Figure 2. In general, the primary goal is to establish some integration function I . By combining stimulus variables in factorial design, a joint test of the integration function and the response scale can be obtained. A successful test, therefore, supports not only the integration function but also the working assumption that the response was on an interval scale.

Of course, this approach rests heavily on the empirical validity of simple algebraic integration functions. Functional measurement is not a routine scaling algorithm, and there have been some clear failures of the additive model, most notably in person perception (Anderson, 1974d, Section 2) and in fundamental work on utility theory by Shanteau (1974, 1975). In person perception, fortunately, it has been possible to establish an averaging model in place of the adding model, and in general the evidence shows that numerous integration tasks do obey a simple cognitive algebra.

Rating scales are, of course, subject to various biases, and certain experimental precautions are necessary to get an interval scale. This matter was discussed at the end of the first part of this article, but the distribution effect deserves a separate re-theoretically be fractionated out by use of the range-frequency model, doing so would be unduly cumbersome in many experiments. In fact, nearly all the work on information integration theory has relied on experimental precautions to eliminate distribution effects.

If distribution effects were present, they would typically produce a specific pattern of nonparallelism. Factorial design tends to produce stimulus combinations that have a normal-like distribution of total magnitude, with many combinations of intermediate value and progressively fewer at either extreme. For such distributions, the equal-frequency tendency would shift the judgment of all combinations outward from the center. Stimuli on opposite sides of the center would

shift in opposite directions, and that would produce systematic discrepancies from parallelism. However, this pattern of discrepancy is seldom found, which indicates that distribution effects can be largely eliminated by appropriate experimental precautions. Equally, it indicates that the rating method can yield interval scales.

Magnitude Estimation

Functional measurement is basically neutral in the controversy between ratings and magnitude estimation (Anderson, 1970b, p. 166). In functional measurement, the basic entity is the psychological law, that is, the integration function of Figure 2. That constitutes a substantive criterion for assessing the validity of any proposed response measure. Ratings and magnitude estimation have equal opportunity to meet this criterion of consistency. Ratings have done well in numerous tests of algebraic integration models (Anderson, 1974a, 1974b, 1974d), whereas magnitude estimation has done poorly.

The first evidence on this point was obtained by Stevens himself: "Use various standards," advised Stevens, "for it is risky to decide the form of a magnitude function on the basis of data obtained with only one standard" (1956, p. 6). But Stevens' own Figure 4 shows that different standards yielded markedly different results. Subsequent work has confirmed Stevens' discrepancy and has shown that many other factors also influence the exponent of a power function fit to the data (see Jones, 1974; Marks, 1974; Poulton, 1968). These effects are just as much undesirable biases as the distribution effects for which Stevens criticized rating scales. They bias the exponents, and no satisfactory way has been given to decide which is the true, unbiased value, if one indeed exists.

Other consistency criteria may also be applied to evaluate the method of magnitude estimation. In particular, numerous attempts have been made to establish algebraic models using magnitude estimation as a response measure. Much of this work has been summarized elsewhere (e.g., Anderson, 1974a) and will not be repeated here. Be-

sides the work of the present author and his associates (e.g., Weiss, 1972, 1975), the work of the Scandinavians (e.g., Ekman, 1958; Goude, 1962; Sjöberg, 1966, 1971) and the Oregonians (e.g., Beck & Shaw, 1968; Curtis, Attneave, & Harrington, 1968; Fagot & Stewart, 1969) deserves special mention for their systematic attacks on the problem.

Overall, magnitude estimation has done poorly in these tests. That has led Stevens (1971, 1974) to argue that the experimental task itself imposes a bias. This ad hoc argument is not very convincing, however, in view of the success of the rating method in many of these same tasks.

A different resolution has been suggested by Marks (1974), who has argued that ratings and magnitude estimation are both valid but measure different perceptual qualities. Ratings are seen as measures of sensory dissimilarity, whereas magnitude estimations are seen as measures of sensory intensity. In Mark's view, then, one set of consistency criteria would apply to rating scales, another set to magnitude estimation.

This interpretation, which is akin to that of Treisman (1964) and Eisler (1965), is reasonable and deserves consideration. However, there are at least two grounds for skepticism. Since magnitude estimations do not generally satisfy the simple models for adding and averaging tasks, whereas ratings do, Marks' position makes it necessary to argue that these tasks themselves, regardless of the response measure, cause the subjects to perceive the single stimuli in terms of dissimilarity rather than magnitude. This view is plausible for judgments of differences, but it is hard to see why a set to judge sums or averages should induce a reaction of dissimilarity, nor is it clear how that might operate with more than two stimuli. Moreover, it would seem that perception of magnitude is prerequisite to perception of dissimilarity in at least some of the judgments in question. Having perceived the magnitudes, there is no apparent reason not to add or average them directly.

Of more immediate relevance to Marks' position is whether magnitude estimation

will in fact satisfy consistency criteria. The effects already noted of various contextual and procedural factors on the exponent parameter present a serious inconsistency for which no resolution seems to be in sight. Furthermore, cross-modality matching, which seems to be the main current consistency criterion for magnitude estimation, is not only ambiguous theoretically (Attneave, 1962; Ekman & Sjöberg, 1965; Treisman, 1964; Zinnes, 1969), but also suffers its own inconsistency by yielding different exponents depending on which dimension is used as the response. Stevens and Greenbaum (1966) have applied the name "regression effect" to this inconsistency but that neither resolves it nor even explains it (Kappauf, 1975, p. 407).

It should be emphasized again that the controversy between ratings and magnitude estimation is not basic from the functional measurement point of view. Each response measure is a simple monotone transformation of the other, and either could be used in conjunction with the monotone scaling principle. For the simple algebraic models, possession of an interval response is a convenience but not a necessity.

Nevertheless, Stevens' concern with getting a linear response measure is well taken. Merely as a matter of convenience, interval measures facilitate progress by an order of magnitude compared to rank-order measures. More important, an interval response measure becomes a necessity in many situations, as when the stimuli interact. Stimulus interaction means that the effective stimulus values are not constant, but vary across different stimulus combinations. Unless this interaction itself happens to obey an algebraic model, the simple analyses no longer apply. Although a two-stage approach (Anderson, 1974a, Section II.B.11; 1974b, Section II.D.4) may be feasible in certain cases, the analysis of stimulus interaction will in general be difficult or impossible without a response methodology that can yield an interval scale. For this reason, the development of numerical response methods deserves continuing experimental study.

Distribution Effects as Contrast

Distribution effects are sometimes interpreted to result from perceptual contrast. If the frequency of a given stimulus is increased, then ratings of neighboring stimuli will shift away, seemingly as though their perceptual magnitudes had changed by contrast with the given stimulus. Helson's (1964) theory of adaptation level rests on this interpretation. In Helson's view, the total stimulus distribution sets up an adaptation level that corresponds to the momentary sensory zero, and each single stimulus is perceived relative to that zero. Theoretically, therefore, the distribution effect is a true perceptual effect, analogous to the effect of visual appearance on felt heaviness in the size-weight illusion. Although Helson's view has not been popular in psychophysics, it has had widespread influence in social-personality.

Much of the work on range-frequency theory seems also to reflect a belief that distribution effects are perceptually real (e.g., Parducci & Perrett, 1971, pp. 427, 451; but see also Parducci, 1974, p. 139). Range-frequency theory has never been clear, however, about whether $R(S)$ or $V(S)$ represents perceptual magnitude. Thus, the range-frequency model itself, in contrast to adaptation-level theory, is silent on the question of contrast. The model does assume that $V(S)$ is constant in Equation 1, but that could hold in the percentile form even with a true change in sensory zero.

The standard objection to the contrast interpretation is that the distribution effect is not truly perceptual, but merely linguistic. In this view, the distribution effect reflects a response equalization tendency that arises from the subject's attempt to utilize a semi-arbitrary, relative response language. This long-standing objection (e.g., Campbell, Lewis, & Hunt, 1958; Stevens, 1958) to Helson's interpretation has never been satisfactorily answered. Moreover, sensitive direct tests using sequential dependencies have not found the contrast effect predicted by adaptation-level theory, but instead a very slight assimilation effect (Anderson, 1971, 1972b). Overall, it seems clear that in most

ases the distribution effect is largely if not entirely a nonperceptual language effect (e.g., Anderson, 1974a, Section II.B.8.c; 1974b, Section III.F.6; Upshaw, 1969). The problem is complicated, however, by the fact that true perceptual contrast does seem to appear under certain conditions (see Anderson, 1971, p. 61). Stimulus integration tasks may help clarify this issue. In hetero-modal adding (Anderson, 1974c), for example, it may be feasible to distinguish response equalization of judgments of total magnitude from true perceptual changes induced by varying the distributions of stimuli within the separate dimensions.

Context Effects: Silk Purses from Sows' Ears

Functional measurement provides an important perspective on the study of context effects. In the traditional approach to psychophysical measurement, context effects are ordinarily treated as undesirable biases (Stevens, 1971, p. 448). With functional measurement, however, it becomes possible to utilize context effects as a tool for measurement.

The size-weight experiment is a good example. The visual appearance does not have a sensory connection to felt heaviness in the direct way that the kinesthetic receptors do. Its effect upon the judgment is contextual. To an investigator oriented toward direct measurement of heaviness, the effect of the visual appearance would constitute an undesirable bias to be eliminated experimentally.

In the functional measurement approach, however, this context effect can be put to good use. It has, of course, primary substantive interest of its own. And because it obeys a simple algebraic model, it provides a way to measure the subjective value of gram weight per se. Instead of being an undesirable bias, the context effect provides a base for validated measurement. This example has special significance because the resultant heaviness scale agrees with scales obtained from other tasks that are perceptually quite different, such as weight averaging and cross-modal adding (Anderson, 1970a, 1972a, 1974c). Such cross-task validation

agrees with the emphasis on converging operations by Garner, Hake, and Eriksen (1956). It also suggests an interesting cognitive economy in which the same internal sensory representation appears in perceptually diverse tasks.

Much the same method and moral hold for the contextual distribution effects in Parducci's range-frequency model. These effects have substantive relevance to a general theory of category judgment. But they also can provide a means for getting an interval scale of the perceptual value $V(S)$ (Anderson, 1974a, Birnbaum, 1974; Anderson, Note 1). Here again, a context effect is transformed from an undesirable bias into a working tool.

This perspective on context effects is not novel, for it can be seen in the work of many investigators. The most dramatic example comes from the theory of signal detection (Green & Swets, 1966; see also Anderson, 1974a, Section II.B.9). Here the successful analysis of the Yes-No task depended on reversing the traditional attitude that bias factors should be eliminated, and instead manipulating them experimentally in order to partial them out.

An interesting example with semantic stimuli is given by the positive context effect in person perception (Anderson, 1966, 1974d, Section 4.3). When one trait adjective in a person description is judged, its rating shifts toward the values of the other traits. There has been some controversy whether this positive context effect is to be interpreted as a change in the meaning of the rated trait, or merely as a semantic illusion (Anderson, Note 2). However, since the effect obeys a simple integration model, it may be used as a basis for scaling regardless of its interpretation.

Other illustrations of this perspective on context effects are discussed elsewhere (Anderson, 1974a, 1974d, in press), and a few may be cited here. Birnbaum, Kobernick, and Veit (1974) report a study of numerosity judgments that is important for its account of contextual learning effects. Resle (1971) and his associates have attempted to apply adaptation-level theory to the anal-

ysis of geometrical illusions. The potential of signal detection theory for scaling has been cited above. The work of Price-Williams (1954) and others suggests that interactions of time and distance may also obey simple algebraic models. Such stimulus integration, like that in the size-weight illusion, can be useful as a natural basis for scaling both conscious and preconscious sensation.

Context effects may thus be widely useful. Systematic manipulation of a context effect can provide a base and frame for validated measurement scales. This contextual measurement approach is not a routine scaling procedure, of course, since it depends on obtaining a valid, substantive model for the context effect. In this functional perspective, measurement theory becomes an organic component of substantive theory.

Psychophysical Law and Psychological Law

The psychophysical law is the function that relates the physical stimulus to its subjective counterpart. The psychological law is the function that relates the subjective stimulus values to the covert response (see Figure 2). An investigator may choose to focus on one or the other of these laws. This choice can have far-reaching consequences, in terms of experimental task, of design and data analysis, and of theoretical direction.

This distinction is well illustrated in a difference that may seem merely technical, but is fundamental. The psychophysical law, as ordinarily conceived, is a function of one variable. This one-variable representation is a natural consequence of the direct phenomenological correspondence between physical dimensions and sensory qualities such as loudness and brightness. It leads equally naturally to the study of judgments of single stimuli, using procedures intended to produce a simple, pure stimulus situation in which contextual factors can be ignored.

The psychological law, in contrast, is ordinarily a function of two or more variables. This several-variable representation is a natural consequence of a concern with stimulus integration. Indeed, the psychological law

is the integration function for a multiple stimulus field.

This one-variable-several-variable distinction makes a decisive difference for measurement theory. A basic problem for any theory of measurement is to provide a basis for transforming an ordinal scale into an interval scale. A function of only one variable is inadequate because it provides no constraints on such monotone transformations. But with two variables in a factorial design, it is possible to allow for monotone transformation and still retain degrees of freedom for testing the psychological law (Anderson, 1962, 1970b). Thus, the psychological law is fundamental, for its substantive interest and as the basis for measurement.³

From this standpoint, it can be argued that the one-variable pursuit of the psychophysical law has had ill effects. Not only is the one-variable approach too narrow to support a theory of measurement itself, but it has misdirected inquiry away from the more fundamental psychological law. This one-variable-many-variable distinction was

³ Recent applications of functional measurement methodology have led to an error theory that allows valid tests of algebraic models even with rank-order data. The first method can be illustrated for an adding model tested in a two-factor design. Available computer programs developed in multidimensional scaling (e.g., Kruskal, 1965; Young, 1972) are applied to make the data as additive as possible separately for each replication. Although the deviations from additivity are all intercorrelated within each replication, they are statistically independent across replications. Accordingly, ordinary repeated measurements analysis of variance may be used to test the model-implied null hypothesis of zero interaction. An invalid model would tend to produce a similar pattern of nonadditivity across replications and so yield a significant interaction term. This test has a bias to reject the null hypothesis but this bias would appear to be small.

The second method relies on the use of two-stage models (Anderson, 1974a, Section II.B.11) that incorporate two algebraic operations. One operation is used to transform the response, the other to test the model. In the three-factor adding model, $A + B + C$, for example, the response would be transformed to make $(A + B)$ as additive as possible. The transformed response should then be an interval scale and so can be used directly in the analysis of variance to test the model-implied null hypothesis of no two-way interaction between C and the compound variable $(A + B)$.

the basic issue in the above discussion of context effects. Other aspects of this same issue arise in the four following statements, which represent dominant themes or ideas in current psychophysics.

Sensation is conscious. This idea is implicit in the various methods of direct estimation, since they ask for introspective reports. However, consciousness is a window on only one stage in the stimulus-response chain. In the size-weight illusion, as noted above, the sensory effect of the physical weight is preconscious, not accessible to introspection. The same holds for many of the physical cues that govern judgments of size and distance or of taste.

These examples are typical of many, perhaps most perceptual tasks. They involve stimulus integration, often across different sense modalities at a preconscious stage. Measurement at the conscious stage will provide a direct linear scale of preconscious sensory effects only under special conditions. In general, such scaling will require methods than can penetrate behind the conscious sensation to dissect it into its preconscious components.

The link between the physical stimulus and the conscious sensation can be treated as a single sensory transducer. This idea was explicit in Stevens' (1960) attempts to use the exponent of a power function fit as an invariant characteristic of the sensory transducer, and it perseveres in his later writings (e.g., Stevens, 1974, p. 370). More generally, it seems implicit in the quest for a single rational form for the psychophysical law.

This idea is not reasonable from either a physiological or a psychological view. Physiologically, many successive stages lie between the physical stimulus and the conscious sensation. To treat their combined action as a lumped unit may have practical utility, but such a lump seems unlikely to have much physiological meaning. Psychologically, conscious sensations are in general resultants integrated across different sense modalities, as already noted. Being sensory compounds, they cannot in general be treated as single sensory representatives.

The psychophysical law has a single, simple mathematical form. Perhaps no idea has been more attractive in the history of psychophysics. Fechner's method of cumulating just noticeable differences would probably have attracted minor interest except that Weber's empirical law led to a theoretical logarithmic function. Later workers have often argued that some other simple law should apply, but few have been willing to abandon the hope for simplicity at this level. This hope seems to have rested largely on the feeling that the psychophysical law was somehow fundamental, and by nature deserved to have a simple form. In part also, the one-variable approach had little alternative but to hope for some simple form to appear.

Yet this idea is not credible from either a physiological or a psychological standpoint. The reasoning here is much the same as for the previous idea. Furthermore, the psychophysical function must be expected to differ between senses and across organisms. Log and power functions may be quite useful as rough numerical approximations, but it is unlikely that they will be exact or general. Much of their popularity seems to have resulted from lack of any pragmatic alternative. The many-variable approach provides such an alternative, by not attempting to impose another arbitrary assumption and instead providing a method to detect how nature actually operates.

The psychophysical law has fundamental importance. The sources of this idea are not entirely clear. In part, it seems to stem from Fechner's original concern with the mind-body problem, in part also from the idea, noted above, that the psychophysical law represents a single sensory transducer. And in part, perhaps, it was felt that discovery of the psychophysical law would solve the problem of measurement. Unfortunately, this orientation toward the psychophysical law has led to methodology that is too narrow to resolve its own problems (Anderson, 1970b).

In the functional measurement approach, it is the integration function or psychological law that is basic. It provides the foundation

for measurement, both conceptually and technically. Conceptually, the measurement scales attain their existence and definition by virtue of the integration function. Technically, the integration function provides the frame for scaling the response and also for scaling the stimuli. Measurement is thus cofunctional with the psychological law.

In this approach, the psychophysical law is derivative from the psychological law. The psychophysical law is no longer fundamental, therefore, and it loses much of the interest and significance that have been attached to it. What interest remains may largely reside in viewing the subjective stimulus values not as primitive sensations but instead as the next object of analysis, themselves perhaps the product of an earlier integration stage.

The essential themes of functional measurement can be seen in much previous work. The study of stimulus integration is time-honored in perception, and many earlier investigators have developed various aspects of the functional measurement approach. Even a partial catalog of algebraic models for perception (Anderson, 1974a) shows how numerous such attempts have been. Of historical interest are Thurstone's (1927) theory of comparative judgment (see Anderson, 1970a, p. 163; 1973) and the work on loudness summation by Fletcher and Munson (1933) and on loudness fractionation by Geiger and Firestone (1933). Among present-day writers, special mention should be made of the work of Birnbaum (1974), Curtis (1970), Garner (1954), Garner, Hake, and Eriksen (1956), Krantz (1972, 1974), Marks (1974), Poulton (1968), Ross and DiLollo (1968), Sjöberg (1966), Sternberg (1969), Treisman (1964), and Weiss (1975), whose views overlap in various respects with the present discussion.

Functional measurement does provide a unified treatment of much previous work. Beyond that, it represents a significant advance in matters of technique, and in empirical accomplishment. Conceptually, however, it may best be considered as a reuniting of measurement theory with the common

sense of common empiricism. In this way, measurement theory resumes its organic role in substantive theory with mutual strengthening. It is to be hoped that the island of psychophysics will in similar way redissolve into the ocean of perception.

REFERENCE NOTES

1. Anderson, N. H. *Algebraic models in perception* (Tech. Rep. CHIP 30). La Jolla, Calif.: Center for Human Information Processing, University of California, San Diego, 1972.
2. Anderson, N. H. *The problem of change-of-meaning* (Tech. Rep. CHIP 42). La Jolla, Calif.: Center for Human Information Processing, University of California, San Diego, 1974.

REFERENCES

- Anderson, N. H. On the quantification of Miller's conflict theory. *Psychological Review*, 1962, 69, 400-414.
- Anderson, N. H. Component ratings in impression formation. *Psychonomic Science*, 1966, 6, 279-280.
- Anderson, N. H. Application of a model for numerical response to a probability learning situation. *Journal of Experimental Psychology*, 1969, 80, 19-27.
- Anderson, N. H. Averaging model applied to the size-weight illusion. *Perception & Psychophysics*, 1970, 8, 1-4. (a)
- Anderson, N. H. Functional measurement and psychophysical judgment. *Psychological Review*, 1970, 77, 153-170. (b)
- Anderson, N. H. Test of adaptation-level theory as an explanation of a recency effect in psychophysical integration. *Journal of Experimental Psychology*, 1971, 87, 57-63.
- Anderson, N. H. Cross-task validation of functional measurement. *Perception & Psychophysics*, 1972, 12, 389-395. (a)
- Anderson, N. H. Looking for configurality in clinical judgment. *Psychological Bulletin*, 1972, 78, 93-102. (b)
- Anderson, N. H. Functional measurement of social desirability. *Sociometry*, 1973, 36, 89-98.
- Anderson, N. H. Algebraic models in perception. In E. C. Carterette & M. P. Friedman (Eds.), *Handbook of perception* (Vol. 2). New York: Academic Press, 1974. (a)
- Anderson, N. H. Cognitive algebra. In L. Berkowitz (Ed.), *Advances in experimental social psychology* (Vol. 7). New York: Academic Press, 1974. (b)
- Anderson, N. H. Cross-task validation of functional measurement using judgments of total magnitude. *Journal of Experimental Psychology*, 1974, 102, 226-233. (c)
- Anderson, N. H. Information integration theory: A brief survey. In D. H. Krantz, R. C. Atkinson, R. D. Luce, & P. Suppes (Eds.), *Contemporary developments in mathematical psy-*

- chology (Vol. 2). San Francisco: W. H. Freeman, 1974. (d)
- Anderson, N. H. Integration theory, functional measurement, and the psychophysical law. In B. F. Lomov (Ed.), *Modern trends in psychophysics*. Berlin: Deutscher Verlag der Wissenschaften, in press.
- Attneave, F. Perception and related areas. In S. Koch (Ed.), *Psychology: A study of a science* (Vol. 4). New York: McGraw-Hill, 1962.
- Beck, J., & Shaw, W. A. Discrimination of loudness similarity. *Perception & Psychophysics*, 1968, 3, 105-108.
- Birnbaum, M. H. Using contextual effects to derive psychophysical scales. *Perception & Psychophysics*, 1974, 15, 89-96.
- Birnbaum, M. H., Kobernick, M., & Veit, C. T. Subjective correlation and the size-numerosity illusion. *Journal of Experimental Psychology*, 1974, 102, 537-539.
- Birnbaum, M. H., Parducci, A., & Gifford, R. K. Contextual effects in information integration. *Journal of Experimental Psychology*, 1971, 88, 158-170.
- Campbell, D. T., Lewis, N. A., & Hunt, W. A. Context effects with judgmental language that is absolute, extensive and extra-experimentally anchored. *Journal of Experimental Psychology*, 1958, 55, 220-228.
- Curtis, D. W. Magnitude estimations and category judgments of brightness and brightness intervals: A two-stage interpretation. *Journal of Experimental Psychology*, 1970, 83, 201-208.
- Curtis, D. W., Attneave, F., & Harrington, T. L. A test of a two-stage model of magnitude estimation. *Perception & Psychophysics*, 1968, 3, 25-31.
- Eisler, H. The connection between magnitude and discrimination scales and direct and indirect scaling methods. *Psychometrika*, 1965, 30, 271-289.
- Ekman, G. Two generalized ratio scaling methods. *Journal of Psychology*, 1958, 45, 287-295.
- Ekman, G., & Sjöberg, L. Scaling. *Annual Review of Psychology*, 1965, 16, 451-474.
- Fagot, R. F., & Stewart, M. Tests of product and additive scaling axioms. *Perception & Psychophysics*, 1969, 5, 117-123.
- Fletcher, H., & Munson, W. A. Loudness, its definition, measurement and calculation. *Journal of the Acoustical Society of America*, 1933, 5, 82-108.
- Garner, W. R. A technique and a scale for loudness measurement. *Journal of the Acoustical Society of America*, 1954, 26, 73-88.
- Garner, W. R., Hake, H. W., & Eriksen, C. W. Operationism and the concept of perception. *Psychological Review*, 1956, 63, 149-159.
- Geiger, P. H., & Firestone, F. A. The estimation of fractional loudness. *The Journal of the Acoustical Society of America*, 1933, 5, 25-30.
- Goude, G. *On fundamental measurement in psychology*. Stockholm: Almqvist & Wiksel, 1962.
- Green, D. M., & Swets, J. A. *Signal detection theory and psychophysics*. New York: Wiley, 1966.
- Helson, H. *Adaptation-level theory*. New York: Harper & Row, 1964.
- Isaac, P. D. Linear regression, structural relations, and measurement error. *Psychological Bulletin*, 1970, 74, 213-218.
- Jones, F. N. Overview of psychophysical scaling methods. In E. C. Carterette & M. P. Friedman (Eds.), *Handbook of perception* (Vol. 2). New York: Academic Press, 1974.
- Kappauf, W. E. Regression effect in judgments to determine equal response contours. *Perception & Psychophysics*, 1975, 17, 405-410.
- Krantz, D. H. Measurement structures and psychological laws. *Science*, 1972, 175, 1427-1435.
- Krantz, D. H. Measurement theory and qualitative laws in psychophysics. In D. H. Krantz, R. C. Atkinson, R. D. Luce, & P. Suppes (Eds.), *Contemporary developments in mathematical psychology* (Vol. 2). San Francisco: W. H. Freeman, 1974.
- Kruskal, J. B. Analysis of factorial experiments by estimating monotone transformations of the data. *Journal of the Royal Statistical Society (B)*, 1965, 27, 251-263.
- Mandel, J. *The statistical analysis of experimental data*. New York: Wiley, 1964.
- Marks, L. E. On scales of sensation. *Perception & Psychophysics*, 1974, 16, 358-376.
- Montgomery, H., & Eisler, H. Is an equal interval scale an equal discriminability scale? *Perception & Psychophysics*, 1974, 15, 441-448.
- Parducci, A. Category judgment: A range-frequency model. *Psychological Review*, 1965, 72, 407-418.
- Parducci, A. Contextual effects: A range-frequency analysis. In E. C. Carterette & M. P. Friedman (Eds.), *Handbook of perception* (Vol. 2). New York: Academic Press, 1974.
- Parducci, A., & Perrett, L. F. Category rating scales. *Journal of Experimental Psychology*, 1971, 89, 427-452.
- Pollack, I. Neutralization of stimulus bias in the rating of grays. *Journal of Experimental Psychology*, 1965, 69, 564-578.
- Poulton, E. C. The new psychophysics: Six models for magnitude estimation. *Psychological Bulletin*, 1968, 69, 1-19.
- Price-Williams, D. R. The kappa effect. *Nature*, 1954, 173, 363-364.
- Restle, F. Visual illusions. In H. Helson (Ed.), *Adaptation-level theory*. New York: Academic Press, 1971.
- Ross, J., & DiLollo, V. A vector model for psychophysical judgment. *Journal of Experimental Psychology Monograph*, 1968, 77 (3, Pt. 2).
- Shanteau, J. Component processes in risky decision making. *Journal of Experimental Psychology*, 1974, 103, 680-691.
- Shanteau, J. An information integration analysis of risky decision making. In M. Kaplan & S.

- Schwartz (Eds.), *Human judgment and decision processes*. New York: Academic Press, 1975.
- Sjöberg, L. A method for sensation scaling based on an analogy between perception and judgment. *Perception & Psychophysics*, 1966, 1, 131-136.
- Sjöberg, L. Three models for the analysis of subjective ratios. *Scandinavian Journal of Psychology*, 1971, 12, 217-240.
- Sternberg, S. The discovery of processing stages: Extensions of Donder's method. *Acta Psychologica*, 1969, 30, 276-315.
- Stevens, S. S. The direct estimation of sensory magnitudes—loudness. *American Journal of Psychology*, 1956, 69, 1-25.
- Stevens, S. S. On the psychophysical law. *Psychological Review*, 1957, 64, 153-181.
- Stevens, S. S. Adaptation-level vs. the relativity of judgment. *American Journal of Psychology*, 1958, 71, 633-646.
- Stevens, S. S. The psychophysics of sensory function. *American Scientist*, 1960, 48, 226-253.
- Stevens, S. S. Issues in psychophysical measurement. *Psychological Review*, 1971, 78, 426-450.
- Stevens, S. S. Perceptual magnitude and its measurement. In E. C. Carterette & M. P. Friedman (Eds.), *Handbook of perception* (Vol. 2). New York: Academic Press, 1974.
- Stevens, S. S., & Galanter, E. H. Ratio scales and category scales for a dozen perceptual continua. *Journal of Experimental Psychology*, 1957, 54, 377-411.
- Stevens, S. S., & Greenbaum, H. B. Regression effect in psychophysical judgment. *Perception & Psychophysics*, 1966, 1, 439-446.
- Thurstone, L. L. The method of paired comparisons for social values. *Journal of Abnormal and Social Psychology*, 1927, 21, 384-400.
- Treisman, M. Sensory scaling and the psychophysical law. *Quarterly Journal of Experimental Psychology*, 1964, 16, 11-22.
- Upshaw, H. S. The personal reference scale: An approach to social judgment. In L. Berkowitz (Ed.), *Advances in experimental social psychology* (Vol. 4). New York: Academic Press, 1969.
- Weiss, D. J. Averaging: An empirical validity criterion for magnitude estimation. *Perception & Psychophysics*, 1972, 12, 385-388.
- Weiss, D. J. Quantifying private events: A functional measurement analysis of equisection. *Perception & Psychophysics*, 1975, 351-357.
- Young, F. W. A model for polynomial conjoint analysis algorithms. In R. N. Shepard, A. K. Romney, & S. B. Nerlove (Eds.), *Multi-dimensional scaling* (Vol. 1). New York: Seminar Press, 1972.
- Zinnes, J. L. Scaling. *Annual Review of Psychology*, 1969, 20, 447-478.

(Received March 18, 1975)

STATEMENT OF OWNERSHIP, MANAGEMENT AND CIRCULATION (Act of October 3, 1917; Act of October 3, 1942; Act of October 3, 1964)		
1. TITLE OF PUBLICATION Psychological Review		2. DATE OF FILING 10/1/75
3. FREQUENCY OF ISSUANCE Bimonthly		4. ANNUAL SUBSCRIPTION PRICE Member \$6 Non-Member \$15
5. LOCATION OF HEADQUARTERS OF PUBLICATION (Name, street, city, state, and ZIP code) (If not printed, give full address)		
1200 Seventeenth Street, N.W., Washington, D.C. 20036		
6. LOCATION OF HEADQUARTERS OF PUBLICATION (Name, street, city, state, and ZIP code) (If not printed, give full address)		
1200 Seventeenth Street, N.W., Washington, D.C. 20036		
7. NAMES AND ADDRESSES OF PUBLISHER, EDITOR, AND MANAGING EDITOR		
American Psychological Association; 1200 Seventeenth Street, N.W., Washington, D.C. 20036		
8. NAMES AND ADDRESSES OF PUBLISHER, EDITOR, AND MANAGING EDITOR		
George Mandl; Dept. of Psychology, University of California - San Diego, La Jolla, CA. 92037		
Anita Revlin; APA; 1200 Seventeenth St., N.W., Washington, D.C. 20036		
9. OWNER (If owned by a corporation, its name and address must be stated and also immediately thereunder the names and addresses of stockholders owning or holding 1 percent or more of total amount of stock. If not owned by a corporation, the names and addresses of individual owners must be given. If owned by a partnership or other unincorporated firm, its name and address, as well as that of each individual must be given.)		
NAME		
ADDRESS		
10. KNOWN BONDHOLDERS, MORTGAGEES, AND OTHER SECURITY HOLDERS OWNING OR HOLDING 1 PERCENT OR MORE OF TOTAL AMOUNT OF BONDS, MORTGAGES OR OTHER SECURITIES (If owned by a corporation, its name and address must be stated and also immediately thereunder the names and addresses of stockholders owning or holding 1 percent or more of total amount of stock. If not owned by a corporation, the names and addresses of individual owners must be given. If owned by a partnership or other unincorporated firm, its name and address, as well as that of each individual must be given.)		
NAME		
ADDRESS		
11. FOR OPTIONAL COMPLETION BY PUBLISHERS MAILING AT THE REGULAR RATES. 39 U.S.C. 3626 provides that publishers mailing at the regular rates must submit to the post office at least one of the following: (a) a statement of ownership, (b) a statement of circulation, or (c) a statement of financial condition. This statement of circulation is required for all publishers mailing at the regular rates. It must be submitted to the post office at least once a year. It must be submitted to the post office at least once a year. It must be submitted to the post office at least once a year.		
12. EXTENT AND NATURE OF CIRCULATION		
A. TOTAL NO. COPIES PRINTED (Net Press Run)		AVERAGE NO. COPIES EACH ISSUE DURING PRECEDING 12 MONTHS
B. PAID CIRCULATION		ACTUAL NUMBER OF COPIES OF SINGLE ISSUE PUBLISHED NEAREST TO FILING DATE
1. SALES THROUGH DEALERS AND CARRIERS, STREET VENDORS AND COUNTER SALES		9,815
2. MAIL SUBSCRIPTIONS		0
C. TOTAL PAID CIRCULATION		9,815
D. FREE DISTRIBUTION BY MAIL, CARRIER OR OTHER MEANS		0,345
E. TOTAL DISTRIBUTION (Sum of C and D)		9,815
F. COPIES NOT DISTRIBUTED		0
1. OFFICE USE, LEFT OVER, UNACCOUNTED, SPOILED, AFTER PRINTING		8,707
2. RETURNS FROM NEWS AGENTS		1,108
G. TOTAL (Sum of F 1 & 2) (Should equal net press run shown in A)		9,815
I certify that the statements made by me above are correct and complete.		9,760
SIGNATURE OF EDITOR, PUBLISHER, BUSINESS MANAGER, OR OWNER		
Henry A. Anderson		

THEORETICAL NOTE

Parallel Psychometric Functions from a Set of Independent Detectors

David M. Green
*Laboratory of Psychophysics,
Harvard University*

R. Duncan Luce
*School of Social Sciences,
University of California, Irvine*

Consider combining several elementary detectors by the extreme decision rule of responding "no" only when all elementary detectors respond "no" and "yes" otherwise. The question raised is: Which psychometric functions for the detectors have the property that the resulting psychometric function is simply the original function displaced in the logarithm of the physical scale? The answer is $p(I) = 1 - e^{-\alpha I^\beta}$, $\alpha, \beta > 0$.

Detection data are often assumed to arise by combining the decisions of a number of more elementary detectors. One extreme assumption is that the elementary detectors are binary and the final decision is affirmative if and only if at least one detector reports a detection. Obviously a number of other decision rules are possible—one might, for example, require some fraction of the elementary detectors to report detection. Such rules are more flexible and may prove more prudent or realistic in particular detection situations. Smith and Wilson (1953) have analyzed such systems in some detail. The appeal of the extreme rule is its simplicity, and so it is not surprising that discussions of this extreme rule occur in many different areas as a kind of null model. In certain classical psychophysical experiments, for example, the influence of signal duration or area on detection can be treated in this way, as can complex signals composed of one or more separate parts.

The psychometric functions observed in such experiments exhibit a surprisingly simple relation. Consider, for example, two different signal durations. Presumably the shorter duration involves fewer elementary detectors than does the larger one, but when one mea-

sures the psychometric function at each duration as a function of log intensity, one usually finds that both psychometric functions have essentially the same shape. That is to say, these psychometric functions are nearly parallel, the one simply being the other displaced to the left or right by an amount that is a function of the parameter value being varied—in this example, signal duration.

Such results were interpreted by Green, McKey, and Licklider (1959) as unfavorable to the putative model for the following reason. Let p denote the elementary detection probability, and let p_n denote that of a system involving n independent elementary detectors. Assuming the decision rule stated, it is obvious that

$$p_n = 1 - (1 - p)^n. \quad (1)$$

Since the relation between p_n and p is nonlinear, it was argued that the shape of the psychometric function must change as n varies. Certainly this is true for a variety of simple psychometric functions, and if n is large enough, this change should be easily noticed, even in empirical data.

However, as Quick (in press) has pointed out, the function

$$p(I) = 1 - e^{-\alpha I^\beta}, \quad \alpha > 0, \quad \beta > 0 \quad (2)$$

has the property that p_n in Equation 1 is of the same form provided I is multiplicatively displaced by a function of n . Thus, on a logarithmic scale of intensity, $p(I)$ and $p_n(I)$ are parallel. The argument is simple:

$$\begin{aligned} p_n &= 1 - (1 - p)^n \\ &= 1 - (e^{-\alpha I^\beta})^n \\ &= 1 - e^{-\alpha (n^{1/\beta} I)^\beta}. \end{aligned}$$

So the function given in Equation 2 has the

This work was supported in part by the National Science Foundation, the National Institute of Health, the Guggenheim Foundation, the University of California, Irvine, and St. John's College, Cambridge, England. We would also like to express our appreciation for the help given by Norma Graham, David H. Krantz, J. Robson, Brian Wandell, and John I. Yellott.

Requests for reprints should be sent to R. Duncan Luce, Department of Psychology and Social Relations, Harvard University, Cambridge, Massachusetts 02138.

property that has been observed empirically. The main purpose of this note is to prove that it is the only function with the property that for some function a on the positive integers and some cumulative probability function p ,

$$p_n(I) = p[a(n)I]. \quad (3)$$

Before deriving this result and commenting on it, we digress briefly on two points. The first is the issue of false alarms. If we denote by ϵ the false-alarm rate of an elementary detector, then by Equation 1, the false-alarm rate of the system is $1 - (1 - \epsilon)^n$ which, when ϵ is small, is approximately $n\epsilon$. It is clear that if the system has a large number of detectors, its false-alarm rate will be modest only if the false-alarm value of each detector is extremely low. This is one reason that we view the model as an extreme one.

The second point to observe is that Equation 2 has appeared in the detection literature at least as early as Brindley (1963). He assumed that I is proportional to the number of quanta impinging on an elementary detector, that there are n independent detectors, and that m quanta are required to excite any one of them. He claimed to show—though we believe the argument wrong (see Appendix)—that for large n the process approaches that of Equation 2 with $\beta = m$ and $\alpha = n/m$!

The same mathematical problem has been treated in a wholly different literature, namely, that concerned with the distribution of the largest of n identically distributed random variables. There it is shown that Equation 2 is one of three possible forms for the asymptotic distribution. The key paper is Fisher and Tippett (1928); it is summarized in Gumbel (1958) and Feller (1966). Feller's discussion is considerably more careful concerning the conditions under which the solution to the corresponding extremal problem is unique. Yellott, who has been working on a related problem in choice theory (Note 1), brought this literature to our attention and showed us how to use Feller's results.¹

¹ The evolution of the result reported here is as follows. We first proved the theorem under a restriction of the function $a(n)$. Krantz then suggested that a somewhat weaker hypothesis would do, and we proved the results under his condition. Later, we read the draft of an article by Yellott (Note 1), in which he proved a related result without assuming anything about the function $a(n)$, but assuming that p is strictly increasing and continuous. This brought to our attention the close connection of our problem to that of the statistics

Form of the Psychometric Function

THEOREM. Assume that the psychometric function p is a (cumulative) distribution function on the nonnegative reals (i.e., it is increasing and onto $[0, 1]$) and that Equations 1 and 3 hold, then

$$p(I) = 1 - e^{-\alpha I^\beta}, \quad \alpha > 0, \quad \beta > 0. \quad (2)$$

Proof. From Equations 1 and 3,

$$[1 - p(I)]^n = 1 - p[a(n)I].$$

Let

$$g(x) = 1 - p(1/x), \quad (4)$$

then it is easy to see that g is a distribution function, that g^n is the distribution of the largest of n independent random variables, each with the distribution function g , and that

$$g(x)^n = [1 - p(1/x)]^n \quad (\text{Equation 4})$$

$$= 1 - p_n(1/x) \quad (\text{Equation 1})$$

$$= 1 - p[a(n)/x] \quad (\text{Equation 3})$$

$$= g[x/a(n)]. \quad (\text{Equation 4})$$

So if $\mathbf{X}_n$ is a random variable with distribution g^n , then $\mathbf{X}_n/a(n)$ has the distribution

$$\begin{aligned} P\left[\frac{\mathbf{X}_n}{a(n)} \leq x\right] &= P[\mathbf{X}_n \leq a(n)x] \\ &= g[a(n)x]^n \\ &= g(x). \end{aligned}$$

According to Feller (1966, pp. 270–271), if the distributions of $\mathbf{X}_n/a(n)$ tend to a distribution g not concentrated at 0—in this case, they all have the same distribution g —then

$$g(x) = e^{-\alpha x^\rho}, \quad \alpha > 0, \quad \rho < 0.$$

Thus,

$$\begin{aligned} p(I) &= 1 - g(1/I) \\ &= 1 - e^{-\alpha(1/I)^\rho} \\ &= 1 - e^{-\alpha I^\beta}, \quad \beta = -\rho > 0. \end{aligned}$$

Application

If we perform an experiment involving two values of n , say n_1 and n_2 , and find the two intensities, I_1 and I_2 , that produce the same probability of detection, it follows from

$$p_n = 1 - e^{-\alpha(nI/\beta I)^\beta}$$

of extremes. In discussing these matters with Yellott, the question was raised whether any assumptions on p were needed beyond its being a cumulative distribution function. Shortly after that, Yellott came up with a proof for his problem which, slightly adapted, served to prove the result we give here.

that

$$n_1 I_1^\beta = n_2 I_2^\beta,$$

and so

$$\beta = \frac{\log n_1/n_2}{\log I_2/I_1}.$$

For a variety of cases in vision and audition we know that doubling the duration of the stimulus—presumably rendering $n_1/n_2 = 2$ —is compensated by a stimulus change of from 1.5 to 3 db (that is, I_2/I_1 from $2^{1/2}$ to 2). Thus, we estimate β in the range from 1 to 2.

Note that it is important in estimating β to measure intensity or energy rather than some quantity, such as pressure, which is proportional to $I^{1/2}$. In such cases the estimated value of the exponent will appear to be twice as large since, if $x^2 \propto I$ then $I^\beta \propto x^{2\beta}$.

Two Devices Displaying This Property

Suppose a Poisson device produces counts with an intensity parameter $\mu = \alpha I^\beta$, the device is monitored for an interval of time Δ , and it reports a detection if at least one count occurs within the interval. The probability of no counts is $e^{-\mu\Delta}$, so

$$\begin{aligned} P(\text{detection}) &= 1 - e^{-\mu\Delta} \\ &= 1 - e^{-\Delta\alpha I^\beta}. \end{aligned}$$

Such a device and the given decision rule yield the required psychometric function.

The application of such a detection model is interesting. First, it seems likely that such a device is plausible only for brief signals, since it assumes that a single count is taken as evidence for the occurrence of a signal. Second, if one envisions this process as applying when signal duration is varied, then Δ is probably not varied because it would eventually generate high false-alarm rates. Rather the outputs of a set of independent monitors are combined, and the time-intensity trade is generated. The exact form of that trade depends on β , but it is possible to predict trades that are not equal-energy contours.

As Wandell has pointed out to us, the proof of the theorem suggests a second decision mechanism yielding Equation 2 as the psychometric function. Suppose the system notes which of several channels appears to be active, calculates a statistic for each, and compares the smallest value with a criterion (i.e., selects the largest among the reciprocals of the statistics). If n channels are active and $\mathbf{Y}_n$ is the smallest statistic, and if there are constants $a(n)$ such that $a(n)\mathbf{Y}_n$ has a distribution

that tends to p , then p is of the form of Equation 2.

Comparison with Other Psychometric Functions

On the theoretical side, we have shown that Equation 2 is the only psychometric function having the property that if one of n detectors with that psychometric function triggers a response, the resulting psychometric function is exactly parallel (on a logarithmic axis) to the elementary one. Is this property of any empirical value? Is it possible to determine whether empirical data obey this rule? The answers to such questions depend on what alternative psychometric functions are considered and the number of elementary detectors that one presumes are combined. It is impossible to give a general answer to the question, but the following relations provide some insight into the nature of the problem by showing how to compute the expected deviation for whatever alternative psychometric function one wishes to consider.

Let $p(I)$ be the psychometric function given in Equation 2. Let some other psychometric function be given by $\rho(I)$, and denote the difference between them by $\gamma(I) = p(I) - \rho(I)$. The following will be simplified if we omit the argument I and denote $p(I) = p$, $\rho(I) = \rho$, and $\gamma(I) = \gamma$. Consider

$$\begin{aligned} p_n &= 1 - (1 - p)^n = 1 - (1 - \rho - \gamma)^n \\ &= 1 - (1 - \rho)^n + n(1 - \rho)^{n-1}\gamma \\ &\quad - n \frac{(n-1)}{2!} (1 - \rho)^{n-2} \gamma^2 \\ &\quad \dots \text{Binominal expansion,} \\ \rho_n &= 1 - (1 - \rho)^n. \end{aligned}$$

Thus,

$$\begin{aligned} |p_n - \rho_n| &\leq n(1 - \rho)^{n-1} |\gamma| \\ &\quad + \frac{n(n-1)}{2!} (1 - \rho)^{n-2} \gamma^2 \dots \\ &\leq n |\gamma| + \frac{n^2}{2!} |\gamma|^2 + \frac{n^3}{3!} |\gamma|^3 \dots \\ &= e^{n|\gamma|} - 1. \end{aligned}$$

Although the bound may be generous, it does permit an estimate of the error. If one desires the error for the n detector case to satisfy $0 < \epsilon \ll 1$, then it is sufficient that

$$e^{n|\gamma|} - 1 \leq \epsilon,$$

that is,

$$n|\gamma| \leq 1n(1 + \epsilon) \cong \epsilon,$$

and so

$$|\rho - p| = |\gamma| \leq \frac{\epsilon}{n} \quad (5)$$

is sufficient.

A difficulty in applying this result is that one often does not know n ; in fact, it is usually estimated from the data. If a situation can be arranged where n is known, such as n bursts of a signal or n dots in a visual field, then the requirement given by Equation 5 can be tested.

Summary

Given an extreme decision rule, namely respond affirmatively when at least one elementary detector reports a detection, we show that the condition stated in Equation 3 yields Equation 2 as the only possible psychometric function. Some rough bounds on the parameters are given. Two mechanisms that produce such a psychometric function are presented, and the relation between this function and some other psychometric functions is derived.

REFERENCE NOTE

1. Yellot, J. *The relationship between Luce's choice*

axiom, Thurstone's theory of comparative judgment and the double exponential distribution (Social Science Working Paper 74). Irvine: University of California, School of Social Sciences, 1975.

REFERENCES

- Brindley, G. S. The relation of frequency of detection to intensity of stimulus for a system of many independent detectors each of which is stimulated by a m -quantum coincidence. *Journal of Physiology*, 1963, 169, 412-415.
- Feller, W. *An introduction to probability theory and its applications, II*. New York: Wiley, 1966.
- Fisher, R. A., & Tippett, L. H. C. Limiting forms of the frequency distribution of the largest or smallest member of a sample. *Proceedings of the Cambridge Philosophical Society*, 1928, 24, 180-190.
- Green D. M., McKey, J. J., & Licklider, J. C. R. Detection of a pulsed sinusoid in noise as a function of frequency. *Journal of the Acoustical Society of America*, 1959, 31, 1446-1462.
- Gumbel, E. J. *Statistics of extremes*. New York: Columbia University Press, 1958.
- Quick, R. E. A vector-magnitude model of contrast detection. *Kybernetik*, in press.
- Smith, M., & Wilson, E. A. A model of the auditory threshold and its application to the problem of the multiple observer. *Psychological Monographs*, 1953, 67 (9, Whole No. 359).

APPENDIX

THE FAILURE OF BRINDLEY'S ASYMPTOTIC FORMULA

Brindley (1963) claimed to show that

$$p_n = 1 - \{[1 + I + I^2/2! + \dots + I^{m-1}/(m-1)!]e^{-I}\}^n$$

approaches, presumably uniformly,

$$1 - [e^{-I^m/m!}]^n$$

as n becomes large. This is true for $m = 1$, since both equations reduce to $1 - (e^{-I})^n$. It is also true for fixed m and large I since $p_n \rightarrow 1$, but of more practical interest is the convergence for small values of I , and hence measurable p_n . Keeping only first-order terms in I , we need only compare the terms within the brackets of the two equations, for

the latter,

$$e^{-I^m/m!} = 1 - I^m/m!,$$

and for the former,

$$[1 + I + I^2/2! + \dots + I^{m-1}/(m-1)!]e^{-I} \cong 1 - I^2.$$

Raising these to the n th power and again retaining first-order terms yields, respectively,

$$1 - nI^m/m! \quad \text{and} \quad 1 - nI^2,$$

and these are simply not the same growth in I , even for $m = 2$; the divergence is worse for $m = 3$ or greater.

(Received April 15, 1975)

Thesaurus of Psychological Index Terms

Contains over 4,000 terms representing an authoritative, unambiguous, and hierarchically structured vocabulary used by psychology and related disciplines. This vocabulary will facilitate precise content representation of concepts described in psychological literature and allow for more flexibility in automated search and retrieval operations. These two factors promote more effective communication of psychological concepts in original works and research, ultimately aiding the content of publications to be focused in accordance with the interest patterns of psychologists.

This **Thesaurus**, developed by the Psychological Abstracts Information Services, is divided into three sections:

The **relationship section** shows the relationships among concepts that are associated with each other. Term relationships noted are: synonymous, broader, narrower, or related.

The **alphabetical section** presents each preferred term in alphabetical order, enabling rapid identification of terms used for indexing.

The **hierarchical section** contains 17 broad content areas in which terms are presented in descending order according to breadth of concept represented. The 17 categories correspond to major classifications used in *Psychological Abstracts*.

The relationship, alphabetical, and hierarchical sections will assist the user in selecting appropriate terms for different uses.

Administrators

- determine areas of research for an organization
- assure representative expertise in personnel selection

Researchers and Authors

- find synonyms for a term representing a concept
- select terms that represent a concept
- define subject matter topics for a paper or book

Educators

- structure a curriculum
- use as a tutorial aid

Consultants

- search the literature
- aid in retrieval of information

Librarians

- provide access to psychological literature
- aid in cataloging psychology material
- aid in ordering new psychological literature

Students

- locate concepts relevant to psychology
- focus areas of study

The **Thesaurus** contains 350 pages with a hard-cover and costs \$12. Please note that all orders amounting to \$15 or less must be accompanied by full payment except for those orders submitted on institutional purchase order forms.

American Psychological Association

Publication Sales Department
1200 17th Street, N.W.
Washington, D.C. 20036





PUBLICATION MANUAL OF THE AMERICAN PSYCHOLOGICAL ASSOCIATION

Second edition, 1974

Current guide to the preparation of manuscripts
for APA and related journals.
This edition adds many examples, including an entire
sample paper. New sections on student papers,
metrication, and mathematical copy. Enlarged sections
on content and organization of an article,
typing the manuscript, and
all aspects of APA style, including reference
lists and tables. 136 pp.
\$4.50.

Please note that payment must accompany all orders
amounting to \$15 or less except for those
orders submitted on institutional purchase order forms.

**American
Psychological
Association**
Publications Sales Dept.
1200 17th Street, N.W.
Washington, D.C. 20036

MADE BY THE DISABLED
REHABILITATION - INDIA

